Journée restitution des travaux de thèses du programme

FUTUR & RUPTURES

15 février 2018

Posters

Institut Mines-Télécom

Journée Futur & Ruptures

Jeudi 15 février 2018

Posters

	Ecole	Titre de la thèse	Encadrant	Thésard
1	Eurecom	Soft Cache Hits: Alternative Content Recommendations	Thrasyvoulos Spyropoulos	Theodore Giannakas
2	IMT Atlantique	Incremental learning of affordances using interactive and strategical algorithms	Maï Nguyen	Alexandre Manoury
3	IMT Atlantique	Conception d'un nez électronique pour la détection de pathologies Application à la détection d'insuffisance rénale	Cyril Lahuec Laurent Dupont	Paul Le Maout
4	IMT Atlantique	Emergence of Long Term Associative Memories in Recurrent Hebbian Networks Under Noise	Charlotte Langlais	Eliott Coyac
5	IMT Lille Douai	Understanding the affective and stylistic human motion From the definition of style to its recognition Stakeholders	Hazem Wannous - Jean-Philippe Vandeborre	Sarah Ribet
6	Télécom PariTech	Système de Localisation 3D Indoor par Radar Multistatique UWB	Jean-Christophe Cousin - Nel Samama	Nour Awarkeh
7	Télécom ParisTech	Cooperative Communications in very large cellular Networks. Nearest Neighbour Cooperation	Giovanidis Anastasios	Luis David Álvarez Corrales
8	Télécom ParisTech	Gestion de ressources photoniques pour l'application aux réseaux de communications quantique	Isabelle Zaquine	Julien Trapateau
9	Télécom ParisTech	Trust based secure routing for the Internet of Things	Anis Laouiti	Asma Lahbib
10	Télécom SudParis	Towards Testing and Verification in Software Defined Networks	Djamal Zeghlache - Natalia Kushik	Asma Beriri
11	Télécom Ecole de Management	Personal data and regulation Empirical and experimental approach	Grazia Cecere	Vincent Lefrère

Incremental learning of affordances using interactive and strategical algorithms

Thesis context

We aim for a robot capable of planning and performing tasks in a real life environment with a limited prior knowledge, by discovering and learning affordances.

- **Problems:** - creating a representation of this real environment from our sensors
  - learning in stochastic, unstructured, high dimensional spaces
- **Approach:** **active exploration** and **long life learning**, in order to learn how to recognize affordances from low level visual features

RB1 robot used during tests

Actors

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

Authors

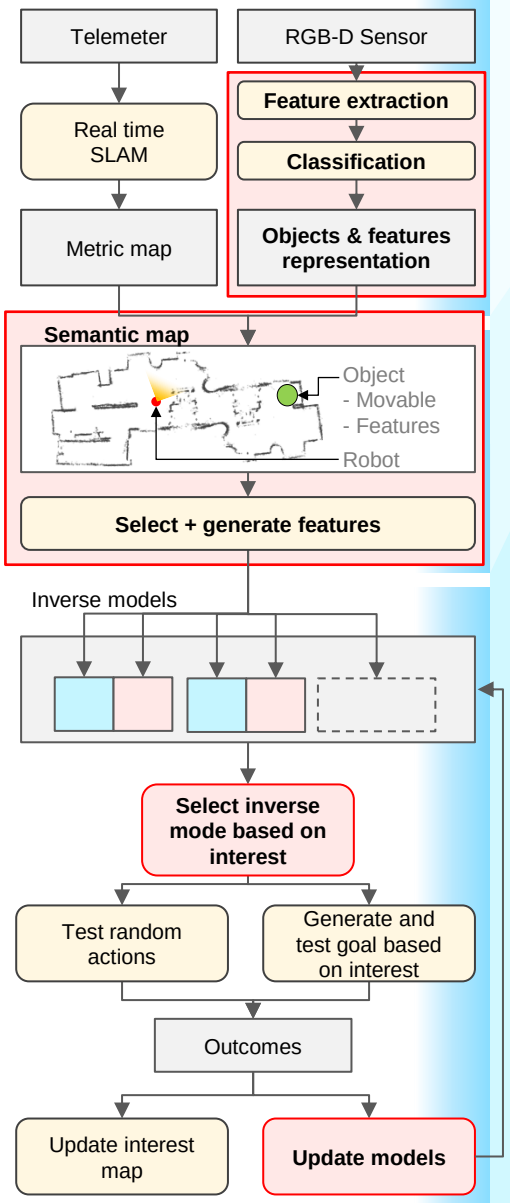
Alexandre Manoury

Thesis supervisor:
Mai Nguyen

Thesis director:
Cédric Buche

Architecture

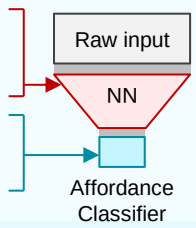
Contributions are marked in red



Visual Extraction

We extract visual features to learn to correlate them with affordances

- **Feature extraction** using a fixed pre trained neural network to segmentate objects and extract their features
- **Affordance classification** using the previously extracted features through a classifier neural network



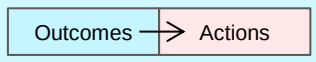
Semantic Map

- Combines **metrics**, **features**, **location**, and **affordances** of objects
- **Generate additional features**: collision features, relative position, …

Learning

Using **Goal babbling** and **Social interaction** *(based on SGIM-ACTS)*

- **Inverse Models** link features and actions



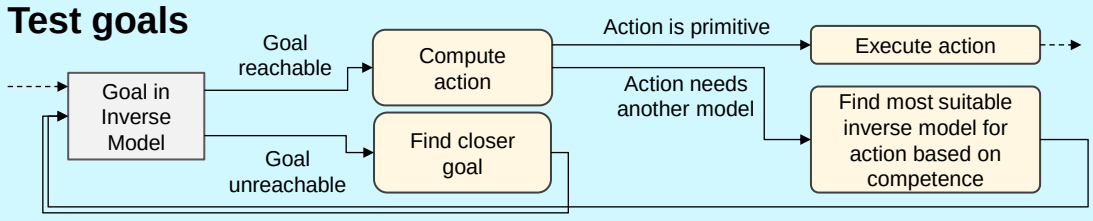
They can be shared across all objects with a given **affordance**

- Update **interest map** based on competence gain or loss

Contributions:

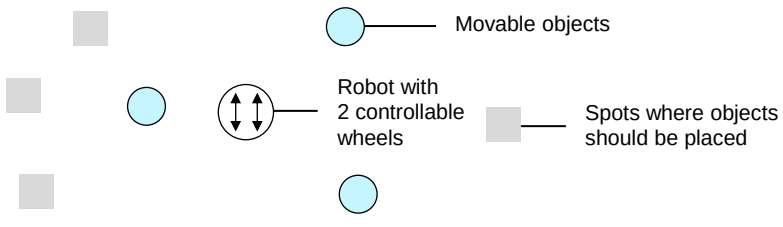
- **Dynamic modelization** updates models depending on the correlation between actions and outcomes
- **Hierarchy** and **planning** are used to reduce the exploration complexity

Test goals



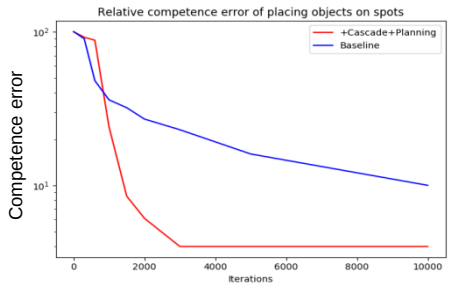
Experiment

- 2D test of the **hierarchy** and **planning** of dynamic models



Results

- Comparison with the algorithm baseline



Models are dynamic in our case and given for the baseline algorithm

Conclusions

- **Hierarchical representation** and **planning** improve learning in high dimensional action and outcome spaces
- **Dynamic models** are able to structure the environment into a hierarchy of inverse models that can be learned

Future works

- Implement the semantic map mechanism
- Test the **complete algorithm** in simulation and in a real experience

Contact : alexandre.manoury@imt-atlantique.fr

15 Février 2018 Journée Futur & Ruptures Paris

Conception d'un nez électronique pour la détection de pathologies

Application à la détection d'insuffisance rénale

Institut Mines-Télécom

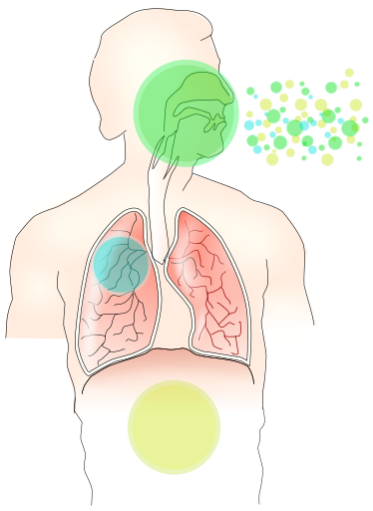
Parties prenantes

IMT Lille Douai
École Mines-Télécom
IMT-Université de Lille

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Auteurs

Paul Le Maout
Cyril Lahuec
Jean-Luc Wojkiewicz
Nathalie Redon
Fabrice Seguin
Laurent Dupont

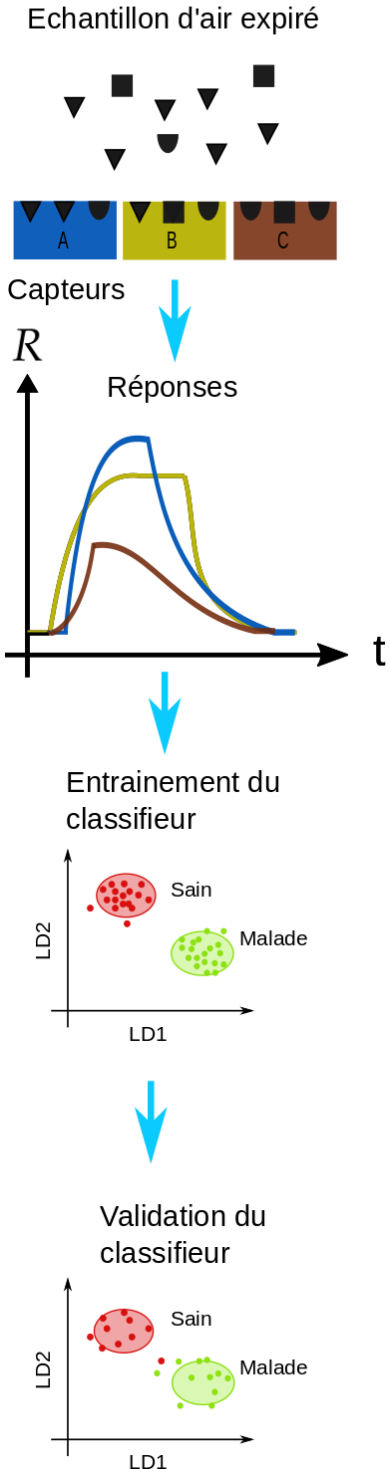


Contexte

- **L'Haleine** est un mélange de plusieurs centaines de composés chimiques (**COVs, ammoniac**) . Une pathologie (cancers, maladies rénales…) modifie la composition de ce mélange.
- Analyser ces mélanges requiert des méthodes spectrométriques lourdes, chères, et compliquées à utiliser.
- L'objectif est de concevoir un **nez électronique** portable capable de discriminer des individus sains des individus atteints de pathologies.

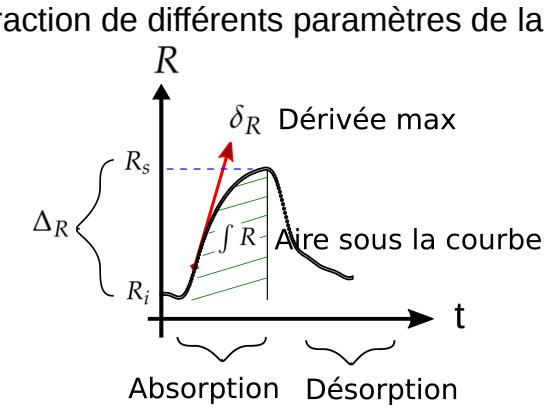
Application à la détection d'ammoniac

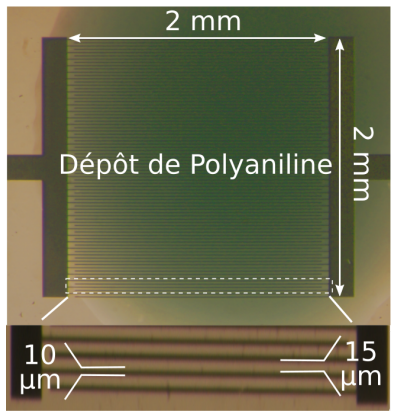
- **Ammoniac** → marqueur intéressant pour les **insuffisances rénales.**
- Matrice de 11 capteurs basés sur un polymère conducteur, la **polyaniline**.
- **Résistance** électrique des capteurs modifiée par l'ammoniac.



Réponses des capteurs et extraction

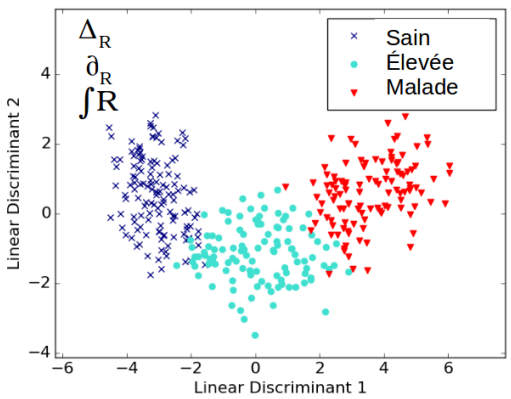
- Extraction de différents paramètres de la courbe.

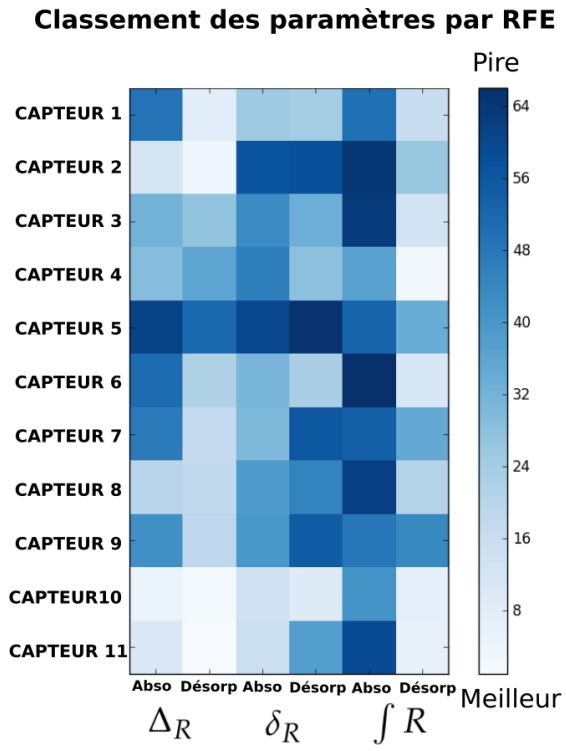




Classification

- Sélection de paramètres + Réduction dimensionnelle + Algorithme de classification





Résultats

- Précision de 91 % avec RFE+LDA+SVM.
- Premier prototype **autonome** réalisé.

Algorithme testé	Δ_R	δ_R	$\int R$	$\Delta_R+\delta_R+\int R$	Sélection avec RFE
LDA	64%	64%	67%	80%	84%
RF	77%	77%	50%	80%	83%
SVM	72%	77%	75%	85%	91%
MLP	75%	78%	63%	85%	87%

Travaux futurs

- Mesures avec la variation de l'impédance complexe.
- Intégration de la partie classification sur circuit numérique (FPGA).
- Mesures avec des échantillons réels (collaboration avec le CHU de Lille).

Journée Futur et Ruptures 2018-PARIS

Contact : paul.lemaout@imt-atlantique.fr

Emergence of Long Term Associative Memories in Recurrent Hebbian Networks Under Noise

Février 2017

Parties prenantes

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Auteurs

Éliott Coyac
Vincent Gripon
Charlotte Langlais
Claude Berrou

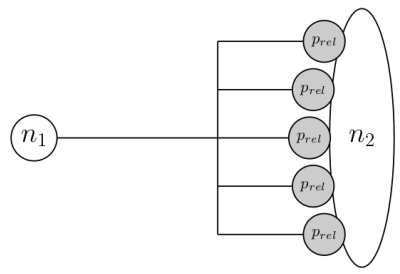
Partenaires

FONDATION Mines-Télécom
La Fondation de l'IMT

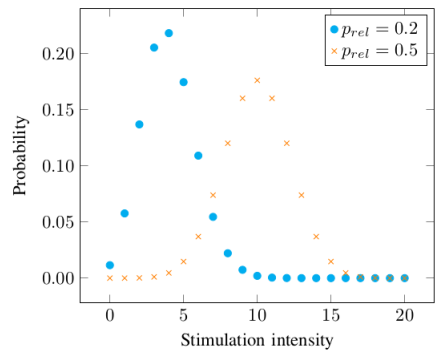
Lab-STICC

A biologically inspired noisy neuron model

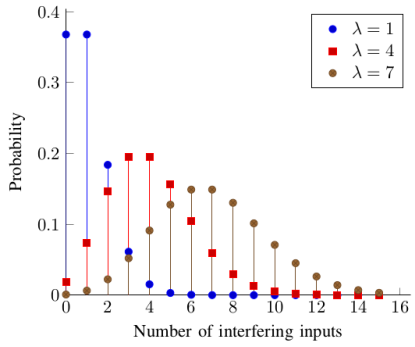
Noise Model



- **Multicanal Synapses** – Synapses are many and have a nonzero probability to fail in releasing neurotransmetters.



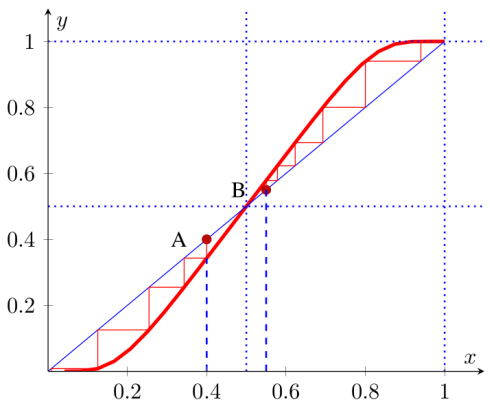
- **Stimulation Intensity** – As a result, stimulation of neurons obeys a binomial law.



- **Interfering Inputs** – We also introduce spontaneous activity from neurons, resulting in possible interferences.

A recurrent clustered neural network model

Network Model



- **Consolidated Hebbian Learning** – The activation function of neurons is chosen so that there are two fixed attractive points: 0 and 1.

$$s(x) \leftarrow \begin{cases} \frac{1}{2} + \frac{1}{2} tanh(tan(\pi x - \frac{\pi}{2})) & \text{if } x \leq 1, \\ 1 & \text{otherwise.} \end{cases}$$

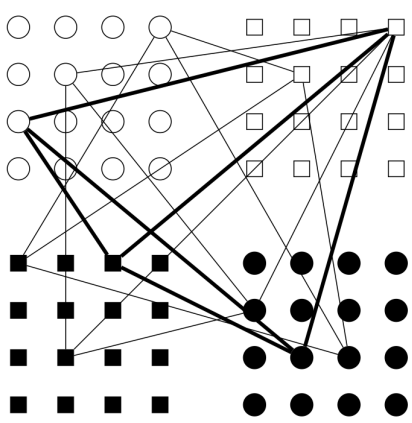
- **Network Equations** – Network equations update at each iteration weigths and activations.

$$h_c(x) \leftarrow \begin{cases} 1 & \text{if } x \in Max_c(X), \\ 0 & \text{otherwise.} \end{cases} \qquad \begin{cases} V(t+1) & = H_c\left(W(t) \cdot V(t) + I(t)\right) \\ W(t+1) & = S\left(\varepsilon \cdot V(t) \otimes V(t) + W(t)\right) \end{cases}$$

- **Properties of the Network** – Locality, boundedness, long-term stability, synaptic depression, incremental learning and competition.

Emergence of long term memory

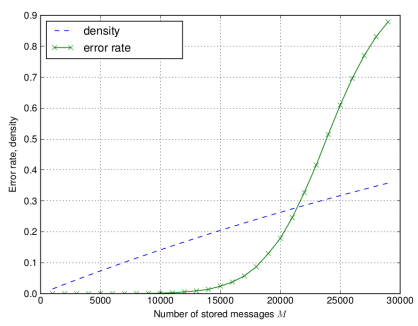
Simulations and Results



- **Neural Clique Networks** – Neural clique networks are autoassociative binary memories built upon clustered neural networks.

- **Performance** – They are able to store then reliably retrieve a lot of messages from noisy inputs, providing density of the network remains small.

- **Emergence of Neural Clique Networks** – We observe that our model is able to make neural clique networks stand out, providing we present each message to store onto the network long enough (number of iterations n_{it} is large enough).



Error rate when retrieving half-erased messages from a network of c = 8 clusters and l = 256 units per cluster.

n_{it}	messages M	connections	added in $\mathcal{N}$	erased in $\mathcal{N}$
50	1000	55888	0	4
60	1000	55634	4	0
50	15000	750774	12	100
60	15000	750922	52	12
70	15000	750570	72	4
100	15000	750926	86	0

TABLE I
CONNECTIONS ADDED AND REMOVED IN NETWORK $\mathcal{N}$ COMPARED TO THE CORRESPONDING NEURAL CLIQUE NETWORK FOR DIFFERENT n_{it}.

Contact: eliott.coyac@imt-atlantique.fr

Institut Mines-Télécom

Understanding the affective and stylistic human motion

From the definition of style to its recognition

Context

- Style brings realism and expressiveness to a motion
- Style not widely studied ≠ action
- Realistic animations obtained from hard work of animators or huge databases
- ➢ Style recognition to enable stylistic motion generation

Stakeholders

IMT Lille Douai
École Mines-Télécom
IMT-Université de Lille

Authors

- Sarah Ribet
- Hazem Wannous
- Jean-Philippe Vandeborre

Partners

CRIStAL
Centre de Recherche en Informatique, Signal et Automatique de Lille

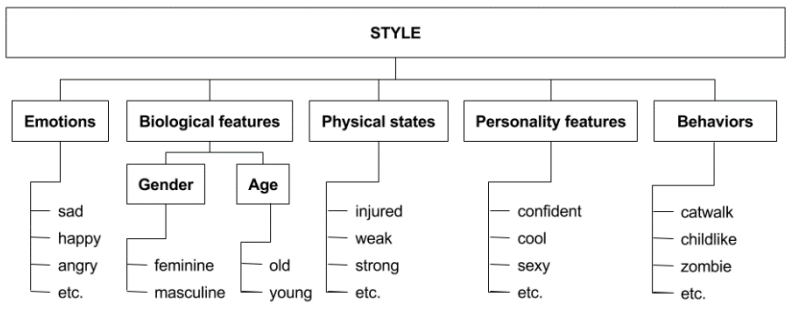
Université de Lille

State of the art and contributions

Style in human body motion

Taxonomy

- Style as a component of motion – added value to the motion action
- Style as variations in a motion – motion type, speed of motion, natural variations of individuals
- Style as individual-related features – behaviors, biological features (age and gender), emotions, personality features, physical state

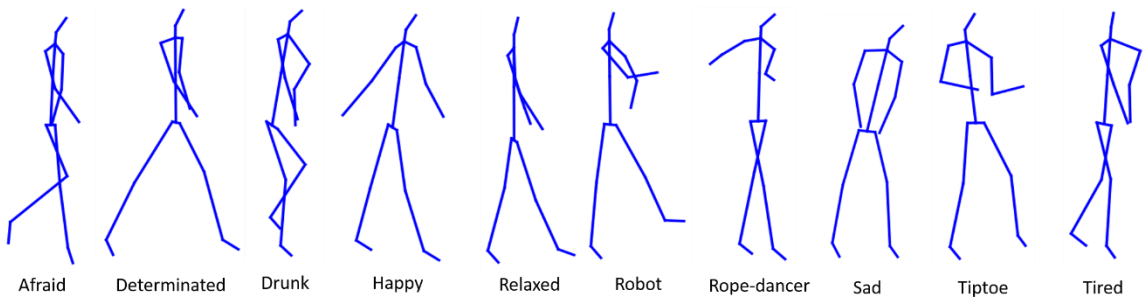


Classification of style when seen as individual-related features

➢ *Spatio-temporal variations of a motion that add value to the motion, depending on individuals*

Data

- 3D skeletal data
- Few consistent data
- ➢ Database creation
  - 1 action: walking
  - 20 subjects ♂/♀
  - 1500 sequences
  - 15 styles: afraid, angry, determinated, drunk, happy, neutral, proud, relaxed, robot, rope-dancer, sad, shy, tiptoe, tired, top model



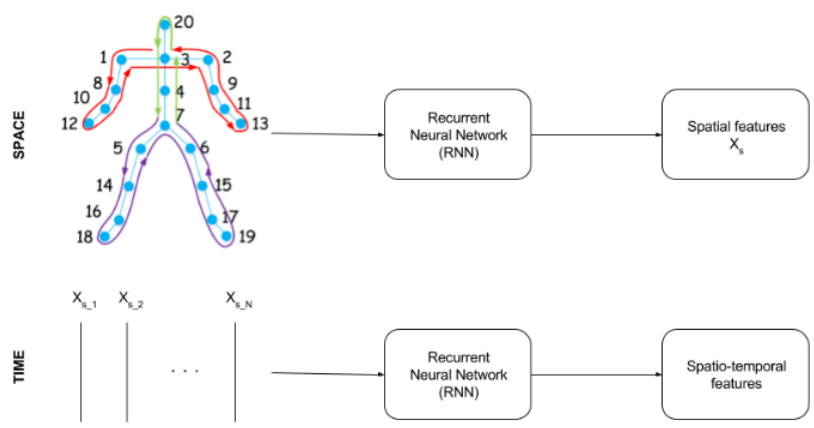
Example postures taken from the created database

First steps to motion style recognition

Recognition pipeline with recurrent neural networks

- Preprocessed sequences of joints as inputs
- Features learned and classified with spatio-temporal Long Short Term Memory (LSTM) networks

Database	Results
MSRAction 3D	~78%
UTKinect	~89%



Features extraction using RNN (LSTM)

Contact : sarah.ribet@imt-lille-douai.fr

February 2018 - Journée restitution thèses Futur & Ruptures Edition 2018

Système de Localisation 3D Indoor par Radar Multistatique UWB

Parties prenantes

TELECOM ParisTech

Auteurs

Nour Awarkeh
Nel Samama
Jean Christophe-Cousin
Muriel Muller

Partenaires

TELECOM SudParis

Contexte et Motivations

- Mauvaise pénétration des signaux GPS à l'intérieur des bâtiments
- Nécessité des informations de localisation pour de nombreuses applications
- Demande croissante d'exactitude dans les systèmes de localisation
- Résolution temporelle très fine du signal UWB

➢ Nécessité de réaliser un système de Localisation 3D Indoor en utilisant la technologie UWB

Modèle de Localisation Indoor

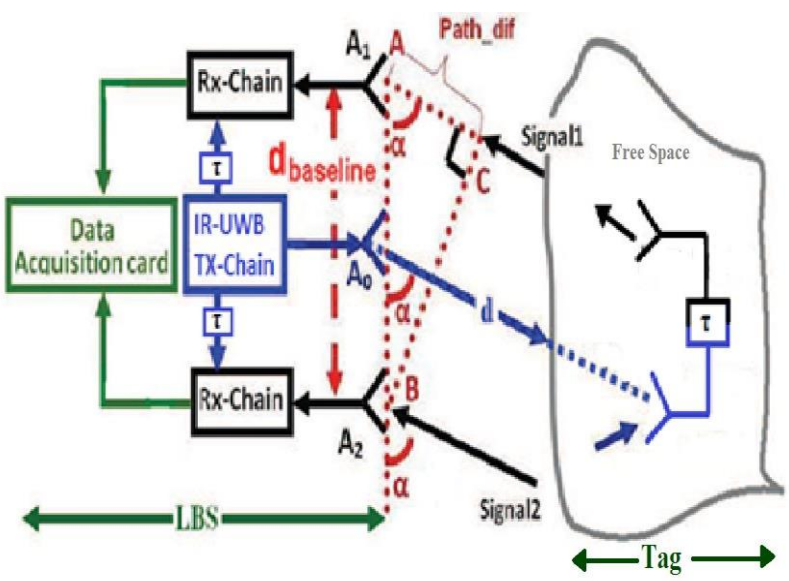


Figure 1. Diagramme du système de localisation

- Impulsion UWB émise par l'antenne A_0 du LBS (station de base de localisation)
- Au niveau du Tag, un retard τ appliqué à l'impulsion pour réduire les phénomènes de rétrodiffusion passive de l'environnement
- Impulsion UWB décalée, reçue par les deux antennes A_1 et A_2
- Corrélation entre les impulsions reçues et l'impulsion de référence décalée du même retard τ
- Traitement numérique et déduction de la distance et de l'angle d'azimut
- En 3D, deux antennes A_3 et A_4 placées perpendiculairement à A_1 et A_2

Validation du Modèle

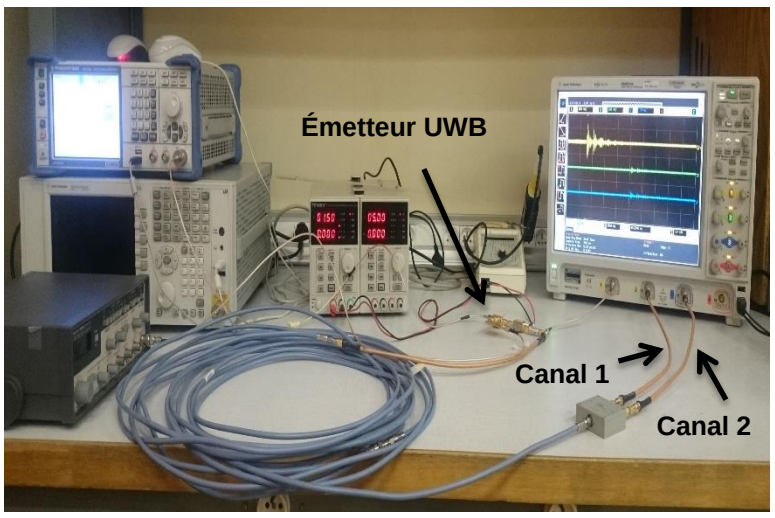


Figure 2. Banc de mesures

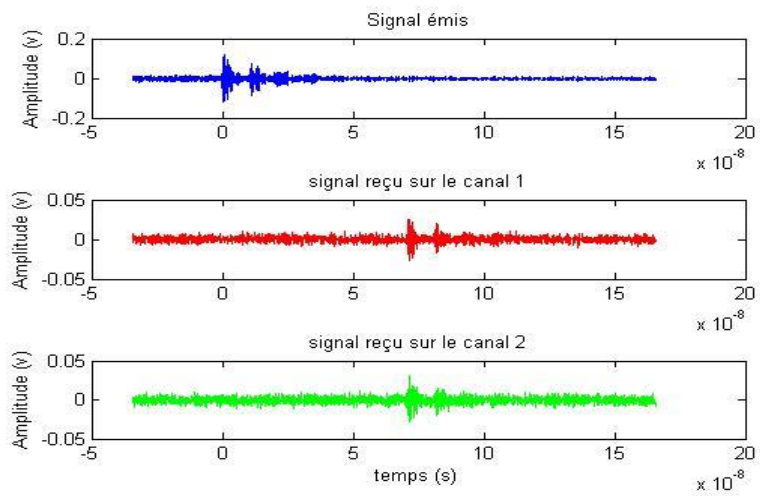


Figure 3. Signal émis et signaux reçus sur les deux canaux

- Validation du modèle réalisée avec des câbles de 50 cm
- Signal émis et signaux reçus sur les deux canaux 1 et 2 présentés dans la figure 3
- Erreur obtenue égale à 0,09 ns qui correspond à une erreur de 3 cm en distance

Contact : nour.awarkeh@telecom-paristech.fr

Février 2018 Colloque de l'Institut Mines Télécom Paris

Cooperative Communications in very large cellular Networks.

Nearest Neighbour Cooperation

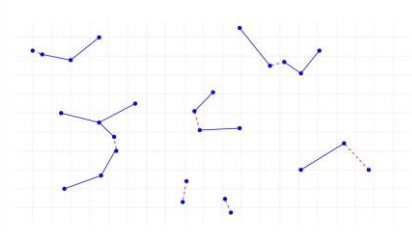
Parties prenantes

Une école de l'IMT

Auteurs

Luis David Alvarez
Anastasios Giovanidis
Philippe Martins
Laurent Decreusefond

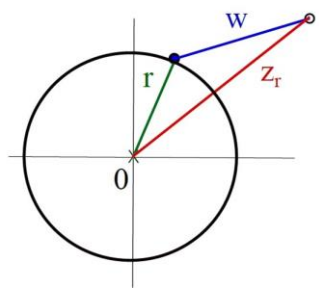
Partenaires



Nearest Neighbour clusters, and each one of their Mutually Nearest Neighbours.

Cooperation in Cellular Networks.

- **Dynamic clusters** – The user chooses the group of Base Stations (BS) for its service. *Problems*: Intensive communication between the BSs, resource sharing.
- **Static Clusters** – The cooperative clusters do not change over time. Problems: The proposed methodologies produce high interference.
- **Static Clusters and proximity** – The static clusters are formed by means of proximity between the nodes: *Random Geometric Graph, Lillypond Model, Nearest Neighbour Model (NNM).*
- **Mutually Nearest Neighbours (MNN)** – Each one of the NN-clusters have only one of these, for which their cooperation is optimal w.r.t. proximity.
- **Modified NNM** – The NN-clusters are recovered, iteratively, from the *root*.

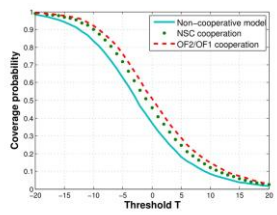


The distance from the typical user to a pair in cooperation is Rice distributed

Properties of the MNNs

When the BSs follow a Poisson Point Process (PPP):

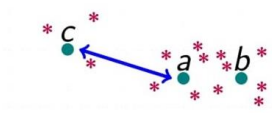
- **38%** of the BSs belong to a cooperative pair and **62%**.
- The distance between MNNs is **Rayleigh distributed**.
- **Attraction** between pairs, **repulsion** among singles.
- **Finite window** analysis and approximation of the **Laplace Transform**.
- It is possible to analyse the benefits of these clusters for a cooperative cellular network through a **PPP-superposition**.
- Considering different **cooperation strategies**, it is possible to get up to a **15%** of **absolute gain**, with respect to the non-cooperative case.



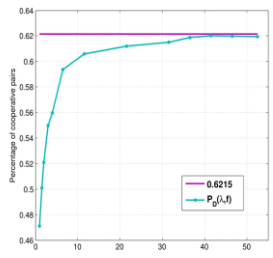
Gains of our cooperative model against the non cooperative case

Resource constraints

- We consider a **new distance** allowing the formation of clusters whose atoms
  - are **geographically close**,
  - They have **enough resources** to make their cooperation beneficial,
  - their amount of resources is **balanced**.
- The **analytic properties** of this distance, along with the assumption that the BSs follow a **marked PPP**, grants an analysis of these cooperative networks.
- For cellular networks whose **resources do not vary a lot**, it is sufficient to analyse them through the original model.
- For cellular networks whose resources **vary a lot**, it is imperative to analyse them via this new clustering methodology.



Problems considering only geographical proximity.



Percentage of pairs in cooperation between both models.

Extensions

- Extend the analysis to **larger classes** of Point Processes.
- **Robust methods** against the inaccuracy of the system's parameters.

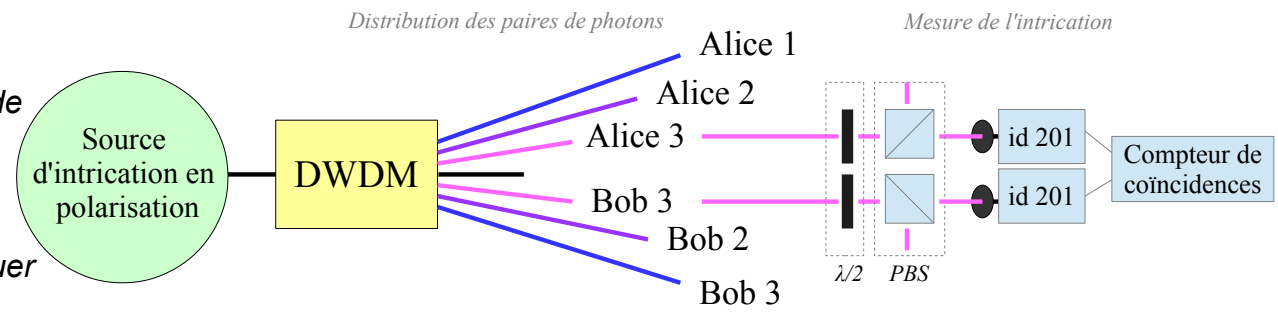
Contact : luisdavidxd@hotmail.com, ananastasios.giovanidis@lip6.fr, philippe.martins@telecom-paristech.fr, laurent.decreusefond@telecom-paristech.fr

Février 2018 Journée Futur & Ruptures Paris

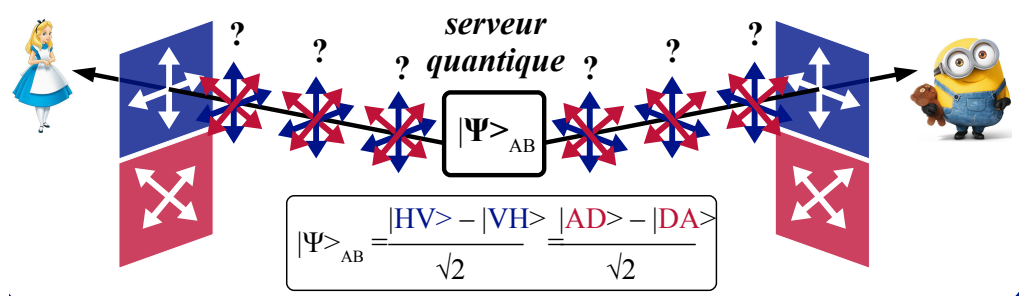
Gestion de ressources photoniques pour l'application aux réseaux de communications quantique

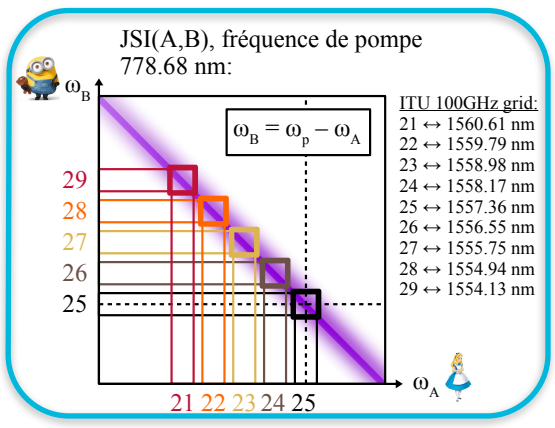
Institut Mines-Télécom

- **Fluorescence paramétrique:** *processus d'optique non linéaire permettant d'obtenir des paires de photons intriqués.*
- **Démultiplexage en longueur d'onde:** *Dispositif des télécommunications pour distribuer les de photons intriqués.*

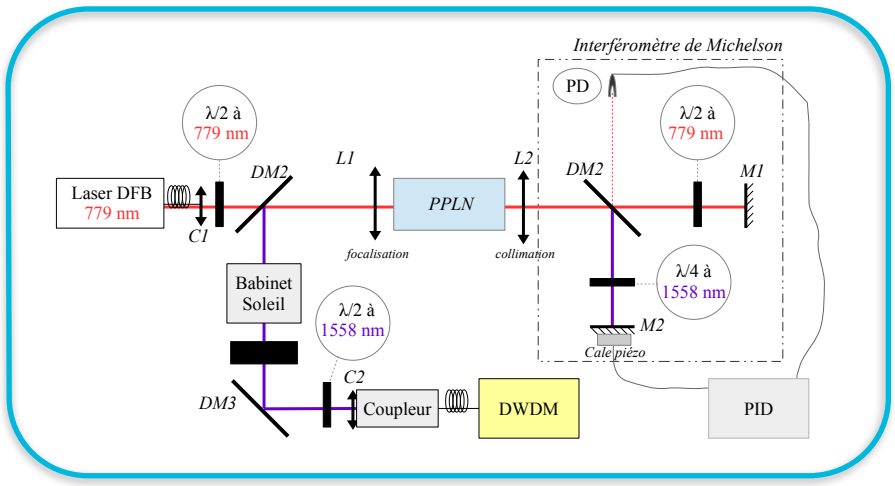


- **Protocole BBM92** : *Protocole de cryptographies utilisant des paires de photons intriqués.*

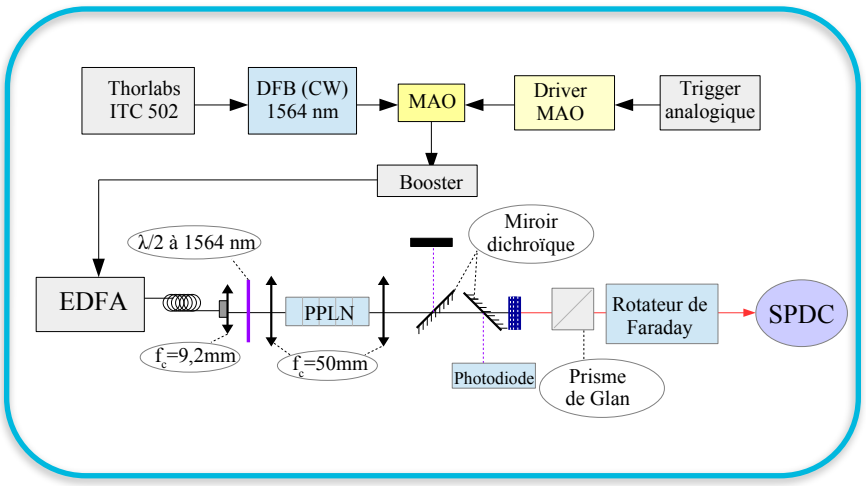




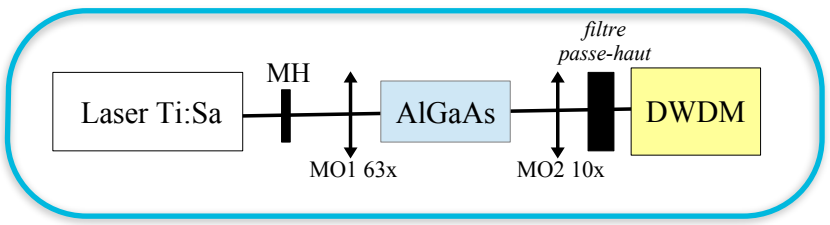
Source continue avec un cristal PPLN [3]



Source pulsée avec un cristal PPLN



Source continue avec un guide d'one AlGaAs [2]



Exemple d'une courbe de visibilité

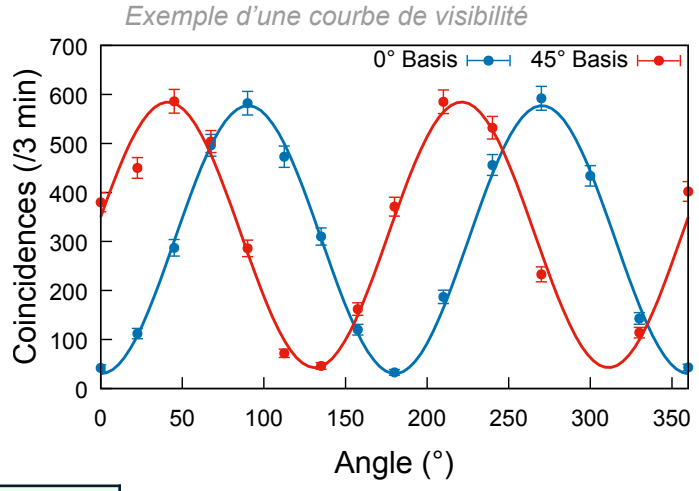


Tableau comparatif

Sources	Canaux	V	QBER	Rsift (bit/s)	Rkey (bit/s)
AlGaAs	21-29	90,0±1,0%	5,0±0,5%	9,91±0,23	3,79±0,10
	22-28	83,5±2,1%	8,2±1,0%	13,0±0,27	3,79±0,10
	23-27	86,7±1%	6,6±0,5%	13,8±0,3	3,28±0,08
	24-26	87,9±1,4%	6,9±0,2%	6,9±0,2	1,95±0,08
PPLN contiune	21-27	85,94±1,8%	7,0±0,8%	19,99±0,26	5,36±0,09
	22-26	87,76±1,3%	6,10±0,6%	18,97±0,28	6,36±0,12
PPLN pulsée	X_X	86,9±1,§%	6,5±0,68%	130,4±1, »6	36,6±0,48

Visibilité

$$V = \frac{\sum C_{max} - \sum C_{min}}{\sum C_{max} + \sum C_{min}}$$

Taux de coïncidence en bases identiques:

$$R_{sift} = \frac{\sum C_{max} - \sum C_{min}}{\tau_{histo}}$$

QBER & taux de clés secrètes [1]:

$$e = \tfrac{1}{2}(1 - V)$$

$$R_{key} \geq R_{sift}(1 - f(e)H_2(e) - H_2(e))$$

Parties prenantes

TELECOM ParisTech

Auteurs

Julien Trapateau
Eleni Diamanti
Isabelle Zaquine

Partenaires

université PARIS-SACLAY

Contact : julien.trapateau@gmail.com

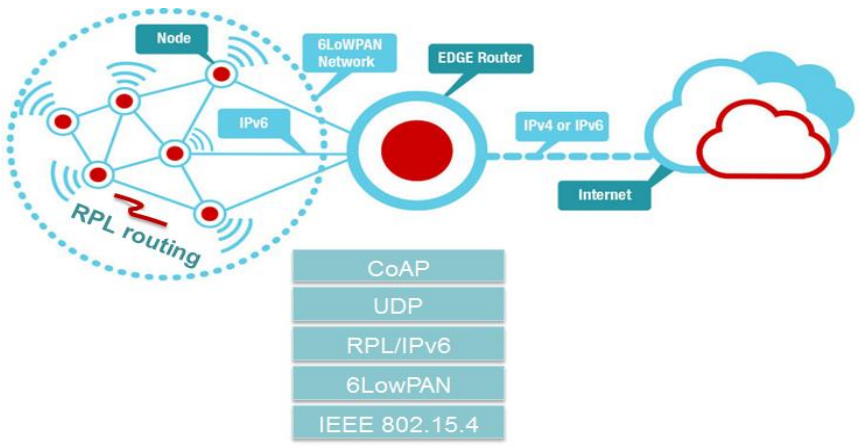
Références:
[1]:X.F. Ma,C-H.F. Fung & H-K. Lo, Phys. Rev. A 76, 012307 (2007)
[2]:C.Autebert et al., Optica 3, 143-146 (2016)
[3]:J.Trapateau, J.Ghalbouni, A.Orieux, E.Diamanti, I.Zaquine Apple. Phys. 118 143106 (2015)

Institut Mines-Télécom

Trust based secure routing for the Internet of Things.

Context

- The Internet of Things (IoT) objects typically collect, communicate and share data that can be used to derive sensitive information and to make decentralized decisions.
- Data packets are transmitted, collected and distributed via RPL, the routing protocol for Low power and Lossy networks considered as the standard protocol of IoT.



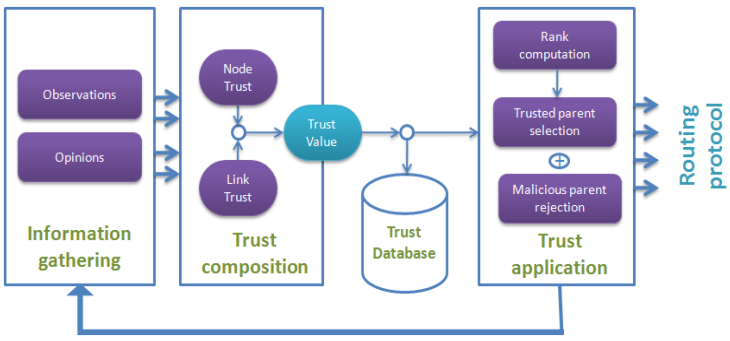
Problematic

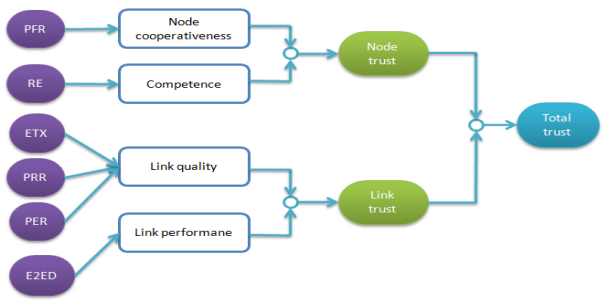
- RPL network is composed of embedded devices with limited power, memory, and processing resources thus their overuse in routing may lead to battery depletion
- The RPL protocol is exposed to a large variety of security attacks causing the loss of a large part of the traffic.
- Participating entities may change their behavior which could disturb the network functioning.

- How to extend the battery life of IoT objects?
- How to be sure that the data received are not corrupted by some malicious nodes in the network?
- How to trust the participating network nodes and how to trust the route data was transmitted over?

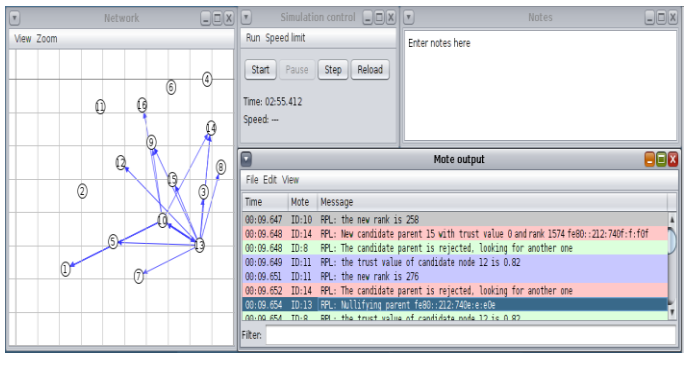
Proposed approach

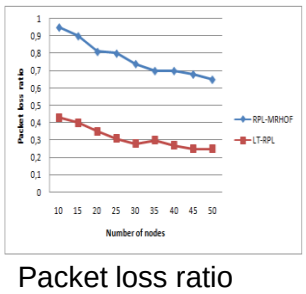
- Enhancing the security aspect of RPL routing protocol.
- Ensuring Trust among entities by considering Trust related to their forwarding behavior as well as that related to the quality of the connecting link.
- Trust computation is based on a set of properties including the reputation parameters, the energy considerations and the QoS factors.
- Integration of the proposed model into the RPL DODAG construction and maintenance phases, Trust values are used for rank computation and thus for preferred parent selection.



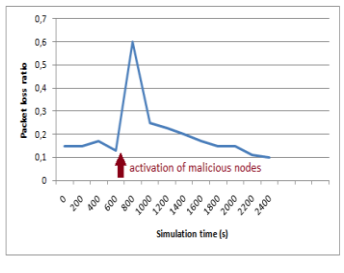


Experiments and results

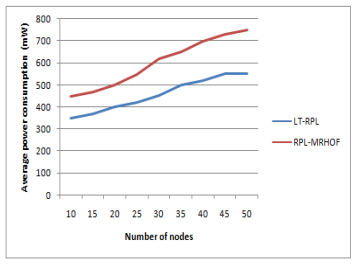




Packet loss ratio comparison between RPL (MRHOF) and LT-RPL under black-hole attacks



Packet loss ratio evolution of LT-RPL in a 50 nodes network size



Influence of the network size on the power consumption

Contact : asma.lahbib@telecom-sudparis.eu

Parties prenantes

TELECOM SudParis

Auteurs

Asma LAHBIB

Partenaires

TELECOM SudParis

samovar UMR 5157

Paris Mines Télécom Futur & Ruptures Février 2018

Towards Testing and Verification in Software Defined Networks

Goal

- ***Formal models* for *testing and verification* of *programmable virtualized networks***
  - *Focus on Software Defined Networking (SDN) components and frameworks*

Motivation

- **SDN adopted to gradually replace traditional networks**
  - Requires testing and validation before deployment
- **Move from concepts to reliable/assured deployment?**
  - SDN systems are complex, multi-component, and heterogeneous
    - Inputs of SDN infrastructures are not traditional, but complex services (e.g. paths and topologies)

- **Existing solutions for guaranteeing correct SDN behaviors are not sufficient**
  - Approaches mostly focused on formal verification and model checking of SDN components or their composition
  - 'Active' testing not largely studied to guarantee correct behavior of SDN infrastructure and its specific components

Novel approaches

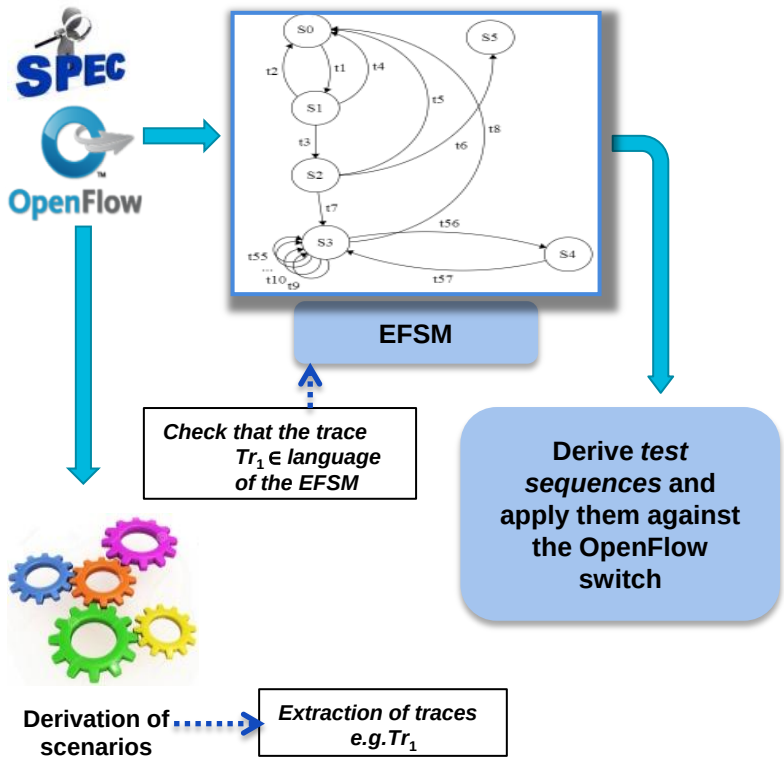
1. Model Based Testing for OpenFlow switch based on Extended Finite State Machine (EFSM) slices

2. Model Based Testing for SDN Architectures based on appropriate *graph enumeration*

EFSM model based testing of OpenFlow switch

Goal: ***Test*** the ***switch*** as a crucial component of an SDN framework

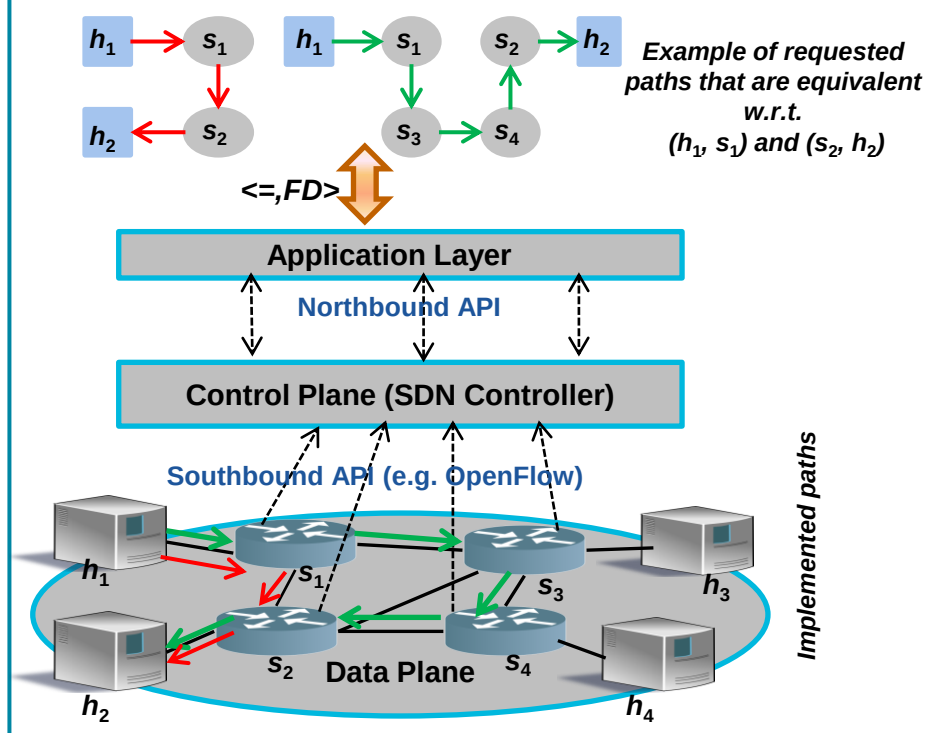
Method: Derive ***test sequences*** that guarantee the desired ***fault coverage*** by formally modeling the switch via an *EFSM*



Testing SDN infrastructures via graph/path enumeration

Goal: ***Test*** the ***entire SDN architecture*** *to check if* ***implemented paths conform to the requested ones?***

Method: Derive ***complete test suites*** w.r.t. the ***fault model <=,FD>*** under ***black box*** and ***white box*** testing assumptions



The set of paths that contains a path of each equivalent class is a complete test suite with respect to the fault model <=,*FD*>

Contributions

- In depth *analysis of the state of the art* indicating the lack of work related to SDN testing
- Novel "*model based testing* technique" for *SDN platforms* relying on appropriate *graph enumeration*
- *Model based* approach for testing the correctness of the *OpenFlow switch*

Future work

- Refine the EFSM model to simplify the test derivation and perform the experiments with OpenFlow switches
- Check the correctness of other SDN components (e.g. compare different controllers)
- Perform experiments on real SDN infrastructures

[1] *A. Berriri, N. Kushik, D. Zeghlache, "On using finite state models for optimizing and testing SDN controller components," Russian Physics Journal, T.59, #8/2. pp. 5-7 , 2016.*

[2] *A. Berriri, J. López, N. Kushik, N. Yevtushenko, D. Zeghlache, "Towards Model Based Testing for Software Defined Networks," Accepted by the 13th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE), 2018.*

Parties prenantes

TELECOM SudParis

INSTITUT Mines-Télécom

Auteurs

Asma Berriri
asma.berriri@telecom-sudparis.eu

Thesis supervised by

Natalia Kushik
natalia.kushik@telecom-sudparis.eu

Djamal Zeghlache
djamal.zeghlache@telecom-sudparis.eu

Personal data and regulation

Empirical and experimental approach

Stakeholder

TELECOM Ecole de Management

Authors

Vincent Lefrere

Grazia Cecere

Fabrice Le Guel

Key figures

Playstore in 2015:
- 1,292,029 free apps
- 85% apps are free

Our Sample:
- 475 867 free apps
- data on privacy
- Data about thirds parties

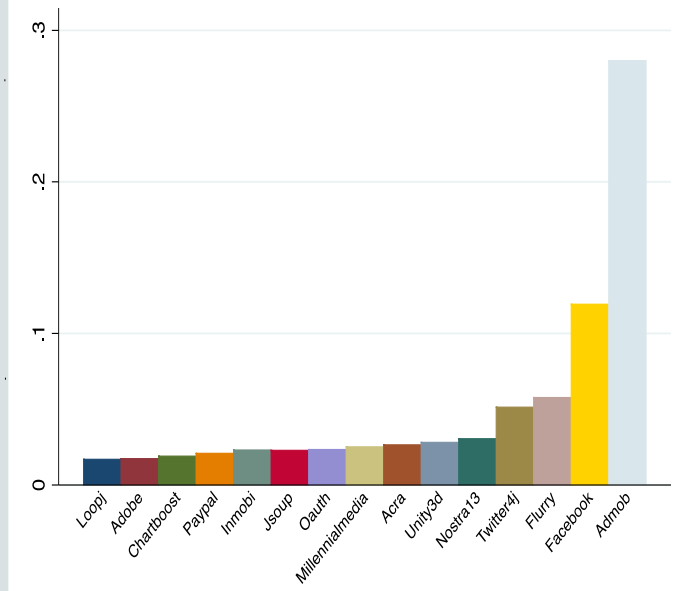
Descriptive results:
- Admob (Google) 86% of market share

Our results:
- Killer apps requested more personal data
- Advertising is for small developers
- Integrated purchase is substitute to the collect of personal data

Research questions and literature:

- In the market for smartphone applications, the majority of apps are zero priced.
- Developers have to monetize theirs apps
- However little is known about their monetization strategies.
- We contribute to three literature strands:
  - Economics of **free digital goods**
  - Economics of **mobile applications**
  - Economics of **privacy**

Strategy of monetization by downloads



Main Objectives:

- Investigate the market of **personal data** through the market of smartphone application
- Improve the understanding of the **free digital goods**
- Explore the **market of smartphone applications**
- Identify the different **strategies of monetization** according the characteristics of the apps
- Analyze the market of **thirds parties** (libraries)

Estimation: Recursive Trivariate probit

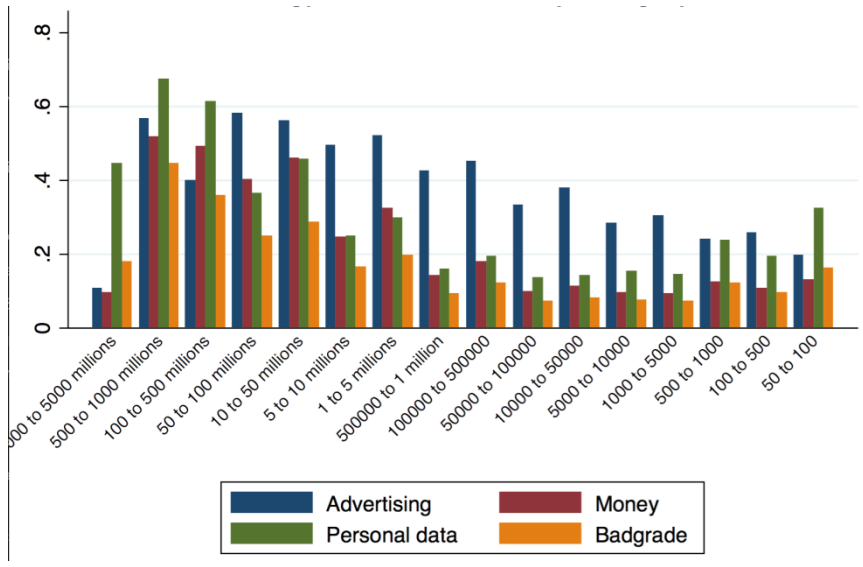
Advertising	Money	Personal data	
Personal data	0.467***	0.626***	
	(0.019)	(0.019)	
Developer website	-0.109***	0.394***	0.189***
	(0.005)	(0.008)	(0.007)
Specialization	-0.467***	0.170***	0.059***
	(0.008)	(0.010)	(0.010)
Privacy policy	-0.122***	0.316***	0.053***
	(0.006)	(0.007)	(0.007)
Killer apps	-0.170**	0.272***	0.907***
	(0.081)	(0.085)	(0.089)
Fixed effects	YES	YES	YES

Python and Android logo

Methodology of data collection:

- First, we download information about apps on the Google PlayStore. We use **python** to scrap apps from the Google **PlayStore**.
- We improve the database by collecting publicly available data on **Privacy Grade**.
- Privacy Grade is an ongoing project of a group of computer science researchers at Carnegie Mellon University.
  - They measure the gap between users' expectations about an app's behavior and the app's actual behavior in **terms of privacy.**

Distribution of top third parties



Conclusion

- Platform can improve transparency in forcing applications to declare the thirds parties to end users
- Within advertising thirds parties, Admob (Google) is the major actor of smartphone application
- There is a link between thirds parties and personal data
- "Big apps" use more personal data compared to the less downloaded apps

Contact : vincent.lefrere@telecom-em.eu
Website: https://sites.google.com/view/vincentlefrere/accueil

15 Février 2018 Journée Futur & Ruptures, Institut Mines Telecom

Institut Mines-Télécom