Journée restitution des travaux de thèses du programme

FUTUR & RUPTURES

15 février 2018

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

Architecture

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

Experiment

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

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

Contact: alexandre.manoury@imt-atlantique.fr
Conception d’un nez électronique pour la détection de pathologies
Application à la détection d’insuffisance rénale

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.

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).

Parties prenantes

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

Contact : paul. lemaout@imt-atlantique.fr
Emergence of Long Term Associative Memories in Recurrent Hebbian Networks Under Noise

A biologically inspired noisy neuron model

**Noise Model**

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

- **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.

**Network Model**

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

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

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

\[
 h_i(x) = \begin{cases} 
 1 & \text{if } x \in \text{Max}_i(X), \\
 0 & \text{otherwise.}
\end{cases}
\]

\[
 \begin{align*}
 V(t+1) &= W(t) \cdot V(t) \cdot f(t) \\
 W(t+1) &= S(z \cdot V(t) \otimes V(t) + W(t))
\end{align*}
\]

- **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 \( N \) is large enough).

<table>
<thead>
<tr>
<th>( n_{xy} )</th>
<th>messages</th>
<th>connections</th>
<th>added in ( N )</th>
<th>erased in ( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1000</td>
<td>55888</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>60</td>
<td>1000</td>
<td>55634</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>15000</td>
<td>750774</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>60</td>
<td>15000</td>
<td>750922</td>
<td>52</td>
<td>12</td>
</tr>
<tr>
<td>70</td>
<td>15000</td>
<td>750570</td>
<td>72</td>
<td>4</td>
</tr>
<tr>
<td>100</td>
<td>15000</td>
<td>750926</td>
<td>86</td>
<td>0</td>
</tr>
</tbody>
</table>

**TABLE I**

Connections added and removed in network \( N \) compared to the corresponding neural clique network for different \( n_{xy} \).
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

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
  - 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 \( \varnothing \varphi \)
    - 1500 sequences
    - 15 styles: afraid, angry, determined, drunk, happy, neutral, proud, relaxed, robot, rope-dancer, sad, shy, tiptoe, tired, top model

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

<table>
<thead>
<tr>
<th>Database</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSRAction 3D</td>
<td>~78%</td>
</tr>
<tr>
<td>UTKinect</td>
<td>~89%</td>
</tr>
</tbody>
</table>

Contact: sarah.ribet@imt-lille-douai.fr
**Système de Localisation 3D Indoor par Radar Multistatique UWB**

**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**

- Impulsion UWB émise par l’antenne A₀ 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 A₁ et A₂
- 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 A₃ et A₄ placées perpendiculairement à A₁ et A₂

**Validation du Modèle**

- 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
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.

**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.

**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.

**Extensions**

► Extend the analysis to larger classes of Point Processes.

► Robust methods against the inaccuracy of the system’s parameters.
Gestion de ressources photoniques pour l’application aux réseaux de communications quantique

- 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’une AlGaAs [2]

Tableau comparatif

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

Références:
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
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

**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


Personal data and regulation
Empirical and experimental approach

Research questions and literature:
► In the market for smartphone applications, the majority of apps are zero priced.
► Developers have to monetize their 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

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.

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 third parties (libraries)

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