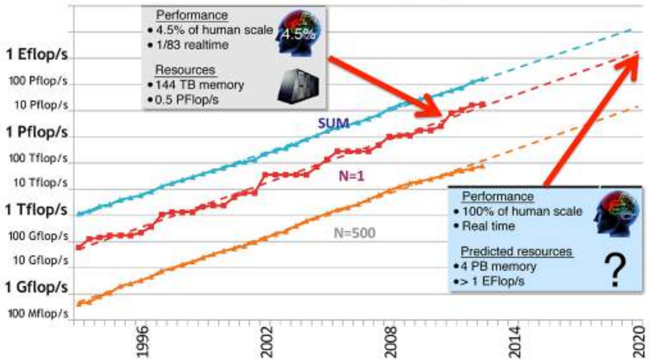
Appréhender l'hétérogénéité à (très) large échelle

Ordonnancement de tâches sur architectures hybrides

La demande en puissance de calcul semble infinie. Afin de dépasser la limite des super-calculateurs actuels (10^{16} Flop/s), des ruptures technologiques architecturales sont nécessaires et rendent les techniques d'ordonnancement obsolètes.

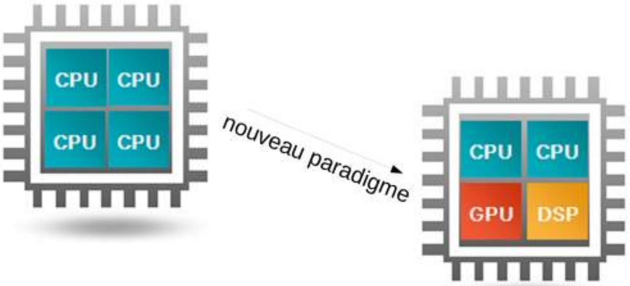
Comment prendre en compte ces différences d'architecture à faible coût ?

Pourquoi exaflopique (10^{18} Flop/s) ?



Évolution de la puissance de calcul & Projection des besoins en ressources

Hétérogène ?



Ce changement de paradigme fait suite à l'évolution des contraintes : applications data-intensive, coût de l'énergie,…

Il permet en effet de mieux maîtriser la consommation énergétique ainsi que les différents types de parallélisme des différents noyaux de calcul.

Avoir des unités de calcul spécialisées permet un gain substantiel de puissance théorique de calcul. Néanmoins l'exploitation à son plein potentiel de ces nouvelles architectures est difficile du fait de la complexité grandissante des architectures.

[Bleuse *et al.* 2014] *Scheduling Data Flow Program in XKaapi: A New Affinity Based Algorithm for Heterogeneous Architectures.* Euro-Par 2014, 560-571, 2014

[Hochbaum, Shmoys 1987] *Using dual approximation algorithmd for scheduling problems: theoretical and practical result.* JACM, 34(1):144-162, 1987

[Stein, Wein 1997] *On the existence of schedules that are near-optimal for both makespan and total weighted completion time.* Operations Research Letter, 21(3):115-122, 1997

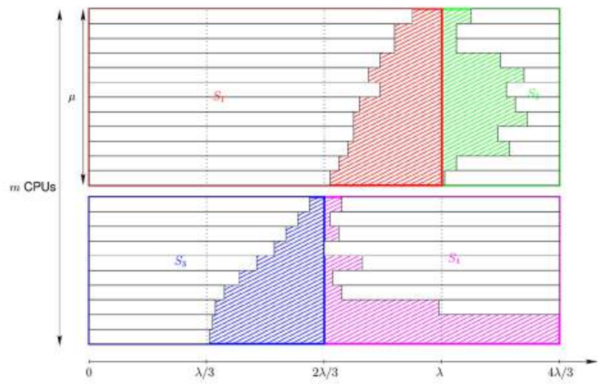
[Turek *et al.* 1992] *Approximate Algorithms Scheduling Parallelizable Tasks.* ACM SPAA'92, 323-332, 1992

Approche proposée [Bleuse *et al.* 2014]

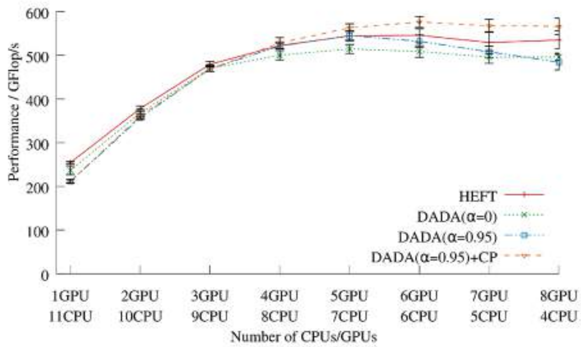
Définition d'une nouvelle abstraction qualitative au travers d'une *affinité contextuelle* (liant unité de calcul et noyau de calcul par ex.).

Combinaison des schémas algorithmiques :

- approximation duale [Hochbaum, Shmoys 1987]
- partitionnement des entrées [Turek *et al.* 1992]
- combinaison de ratio d'approximation [Stein, Wein 1997]



Gain de performances à grain fin



Travaux futurs

Le modèle de tâches moldables permet de contrôler la finesse de parallélisation des noyaux de calcul et d'introduire une notion de compromis temps/espace. Comment interfacer affinité et moldabilité ?

L'affinité permet de prendre en compte de manière qualitative des phénomènes contextuels (communication par ex.). Une telle approche permet-elle des gains de performance à l'échelle de la machine complète ?

Message ?

Une **vision qualitative** de la complexité des nouvelles architectures semble une abstraction raisonnable pour ordonnancer de manière efficace.

Contact : Raphaël Bleuse
raphael.bleuse@imag.fr

Laboratoire d'Informatique de Grenoble 2015

DGA LIG

MOAIS
"Multi-programmation et Ordonnancement sur ressources distribuées pour les Applications Interactives de Simulation"

Sources images :
- Horst Simon, *Why we Need Exascale and Why We Won't Get there by 2020*, Optical Interconnects Conference 2013
- Qualcomm Technologies Inc.

An approach to reasoning about graph transformations

Jon Brenas, Rachid Echahed, Martin Strecker

Reasoning on graph transformations

- It is, in full generality, a hard problem.
- We introduce a kernel language to modify graphs
- We try to make the reasoning as adaptable as possible
- It is often undecidable
- We use a Hoare-style calculus to reason on graph transformations

Programming Language

The programming language we use features:

- fairly usual statements (if statements, while loops, ...) using formulas as conditions
- addition\deletion of a node label: `addC(`i : c`)` `delC(`i : c`)`
- addition\deletion of an edge label: `addR(`$i\ R\ j$`)` `delR(`$i\ R\ j$`)`
- selection: `select` i `with` *cond*

Logic

- We need a logic to define conditions
- During verification, **substitutions** are used. They are:

$R + (i, j)$	(insertion of an edge label)
$R - (i, j)$	(deletion of an edge label)
$c + i$	(insertion of a node label)
$c - i$	(deletion of a node label)

Example program

We give a small example:

R, a researcher, is hired by L, a laboratory:

Pre: $R : Researcher \wedge L : Lab$

The first step of hiring is adding R to the roster:

addR(R Member L);

Then we add all the topics R has written about to the topics of interest of L:

while $(\exists \tau . \neg (L\ \mathtt{T\_i}\ \tau) \wedge R : \exists \mathtt{Pub}.(\exists \mathtt{T}.\{\tau\}))$ {

Inv: $R : Researcher \wedge L : Lab$

select τ **with** $\neg(L\ \mathtt{T\_i}\ \tau) \wedge R : \exists.\mathtt{Pub}(\exists.\mathtt{T}\{\tau\})$;

addR $(L\ \mathtt{T\_i}\ \tau)$

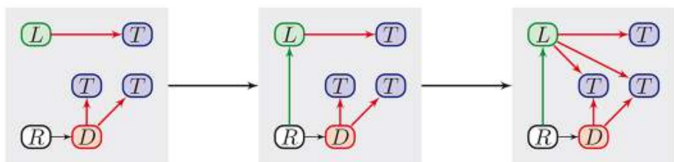
};

In the end:

- R should still be a researcher
- L should still be a laboratory
- each subject of each publication written by R should be a topic of interest of L
- and R should be a member of L

Post: $R : Researcher \wedge L : Lab \wedge R : \forall \mathtt{Pub}.(\forall \mathtt{T}.(\exists \mathtt{T\_i}^{-}.\{L\})) \wedge R\ \mathtt{Member}\ L$

Illustration



Proof obligations

The Hoare-like calculus for the generation of **weakest preconditions**:

$wp(\mathtt{Skip},\ Q) = Q$

$wp(\mathtt{addR}(i_1\ R\ i_2),\ Q) = Q[R := R + (i_1, i_2)]$

$wp(s_1; s_2,\ Q) = wp(s_1, wp(s_2,\ Q))$

$wp(\mathtt{if}\ b\ \mathtt{then}\ s_1\ \mathtt{else}\ s_2,\ Q) = (b \rightarrow wp(s_1, Q)) \wedge (\neg b \rightarrow wp(s_2, Q)))$

$wp(\mathtt{addC}(i\ :\ c)\ Q) = Q[c := c + i]$

$wp(\mathtt{delR}(i_1\ R\ i_2),\ Q) = Q[R := R - (i_1, i_2)]$

$wp(\mathtt{while}\ b\ \mathtt{inv}: f\ s,\ Q) = f$

$wp(\mathtt{delC}(i\ :\ c),\ Q) = Q[c := c - i]$

$wp(\mathtt{select}\ i\ \mathtt{with}\ b,\ Q) = \forall i.(b \rightarrow Q)$

and **verification conditions**:

$vc(\mathtt{if}\ b\ \mathtt{then}\ s_1\ \mathtt{else}\ s_2,\ Q) = vc(s_1, Q) \wedge vc(s_2, Q)$

$vc(\mathtt{select}\ i\ \mathtt{with}\ b,\ Q) = \top$

$vc(s_1; s_2,\ Q) = vc(s_1, wp(s_2, Q)) \wedge vc(s_2, Q)$

$vc(\mathtt{while}\ b\ \mathtt{inv}: f\ s,\ Q) = (f \wedge \neg b \rightarrow Q) \wedge (f \wedge b \rightarrow wp(s, f)) \wedge vc(s, f)$

$vc(\mathtt{addC}(i\ :\ c)\ Q) = \top$

$vc(\mathtt{addR}(i_1\ R\ i_2),\ Q) = \top$

$vc(\mathtt{Skip},\ Q) = \top$

$vc(\mathtt{delC}(i\ :\ c),\ Q) = \top$

$vc(\mathtt{delR}(i_1\ R\ i_2),\ Q) = \top$

Results

- Given a precondition Pre, a list of statements $\mathcal{S}$ and a postcondition $Post$, if $vc(\mathcal{S}, Post) \wedge (Pre \Rightarrow wp(\mathcal{S}, Post))$ is valid then if the initial configuration satisfies Pre and the program stops, the final configuration satisfies $Post$.
- We characterized some sublogics for which the validity of $vc(\mathcal{S}, Post) \wedge (Pre \Rightarrow wp(\mathcal{S}, Post))$ is decidable.

Perspectives

In the future, we aim to:

- add more statements to the programming language
- apply it to real world problems
- find more logics that are stable under substitutions

CAPP "Calculs algorithmes programmes et preuves"

Privacy-Aware Personal Information Discovery

Thiago Moreira da Costa
Supervisor: Hervé Martin

CAPES

what to share? | to whom? | when? | limitless storage | data mining | distributed

Personal Privacy — **Big Data**

How to control personal privacy in the big data era?

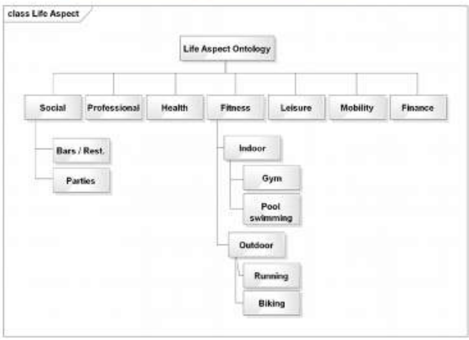
health, fitness, leisure, mobility, professional, social, finance

Individual — **stakeholders**

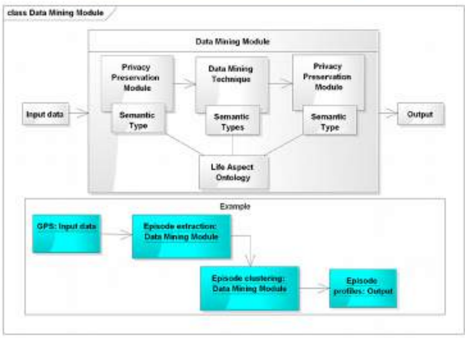
traffic prediction
financial profile
sedentary life style profile
relationship status discovery
health-status discovery
address discovery
addiction discovery
unaware surveillance

Personal Information Discovery Contract

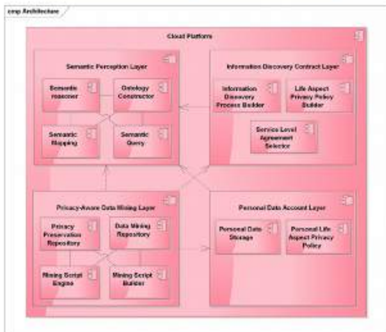
Life Aspect Privacy Policy



Data Mining Process



Privacy-Aware Data Mining Platform



Process

Data Mining and Privacy-Aware Components Registration → Life Aspect Privacy Policy Definition → Information Discovery Contract Definition / Sensor Data and Information Collection → Privacy-aware Personal Information Discovery

STEAMER
"Spatio-temporal information, adaptability, multimedia and knowledge representation"

Multi-scale interaction techniques for the interactive visualization of traces of execution

MOTIVATION

Increasing complexity of MPSoCs

Increasing complexity of embedded software: heterogeneous systems, parallel architecture

Increasing complexity of industry standards (H265, U4K)

Huge impact on cost and time development

EXECUTION TRACES

Video decoding

```
[…]
12.3456 C 1023  4259
12.3458 F 0x538 0x98b
12.3571 f 0x538
[…]
```

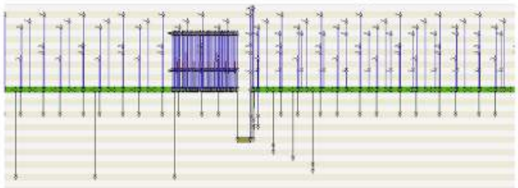
Execution

Trace file

CHALLENGES

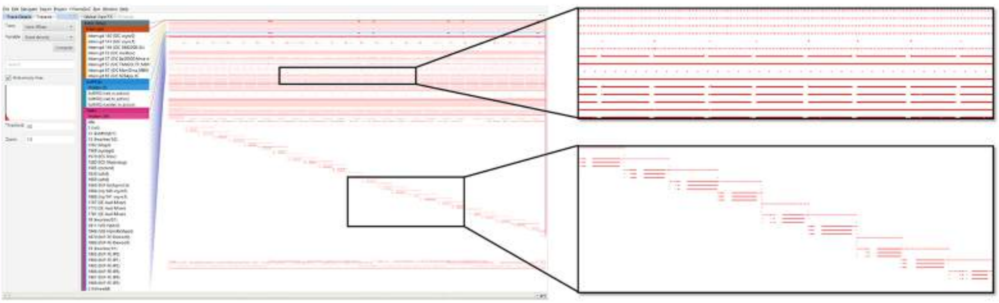
Huge amount of data

Lack of tool to filter and navigate the data



TRACEVIZ

Objectives: Propose a global view of the execution, filter mechanisms, visually detect period and similar behaviors



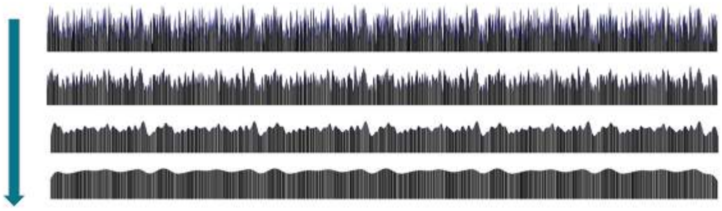
Periodic and similar behavior can be visually detected

Execution patterns appear

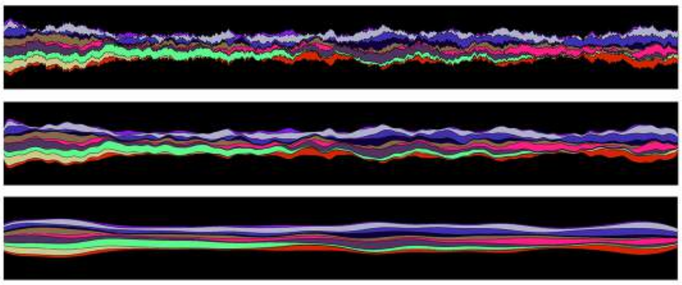
SLICK GRAPH

Objectives: Interactive technique using a Gaussian to smooth the data for an easier reading and period detection

User increases smoothing factor on the data
Shape of the curve is consistent
Periods are easier to read



Smoothing technique is transversal to other time series visualizations. Example with a Stacked Graph



Classical Stacked Graph using the Slick Graph smoothing technique on each of the layers.

IIHM

"Ingénierie de l'Interaction Homme-Machine"

Laboratoire d'Informatique de Grenoble

2015 LIG

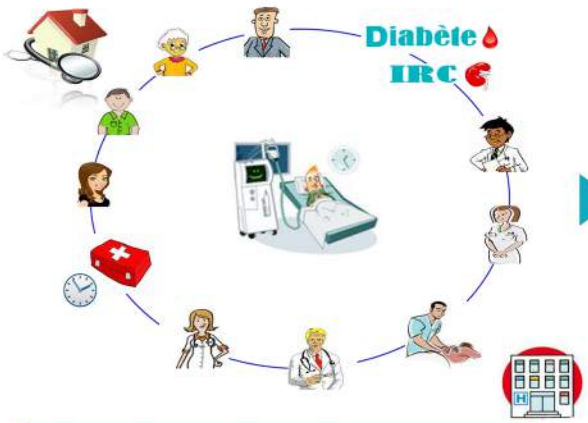
Rémy Dautriche – IIHM/STMicroelectronics

Advisors: Alexandre Termier (Univ. Rennes 1), Renaud Blanch (Univ. Grenoble Alps), Miguel Santana (STMicroelectronics)

Construction de protocoles de soins auto-adaptatifs pour le suivi des maladies

Amira DERRADJI
Thèse Cifre chez Arcan Systems
Directeurs de thèse : Agnès Front, Christine Verdier et Vincent Bouzon

Contexte : Prise en charge de maladies chroniques à domicile



Enjeu sociétal

« Patient expert et éducation thérapeutique » : le vécu quotidien avec la maladie aide le patient à mieux connaître sa maladie et à en tirer une expertise. Il devient donc partie prenante de ses propres soins

Problèmes médicaux actuels

- Protocole de soins non personnalisé pour le patient
- Forme du protocole de soins non adaptée au patient (orale, texte ou graphique simple)
- Suivi du protocole effectué à la main via de simples cahiers de liaison
- Patient non impliqué de manière directe dans les soins
- Pas d'outils de suivi de maladies chroniques, sans (quasi)-intervention médicale

Objectifs de la thèse

- ✓ **Construction et personnalisation** de protocoles de soins
- ✓ **Intégration de l'expertise du patient** (partie prenante de son protocole de soins) en lui permettant de signaler tout imprévu **subi** ou **réalisé**
- ✓ Interprétation de la situation de l'imprévu signalé par le patient et mise en œuvre d'une action ciblée
- ✓ Amélioration de **la connaissance médicale** afin d'aider le professionnel de santé à mieux répondre à de nouvelles situations pathologiques

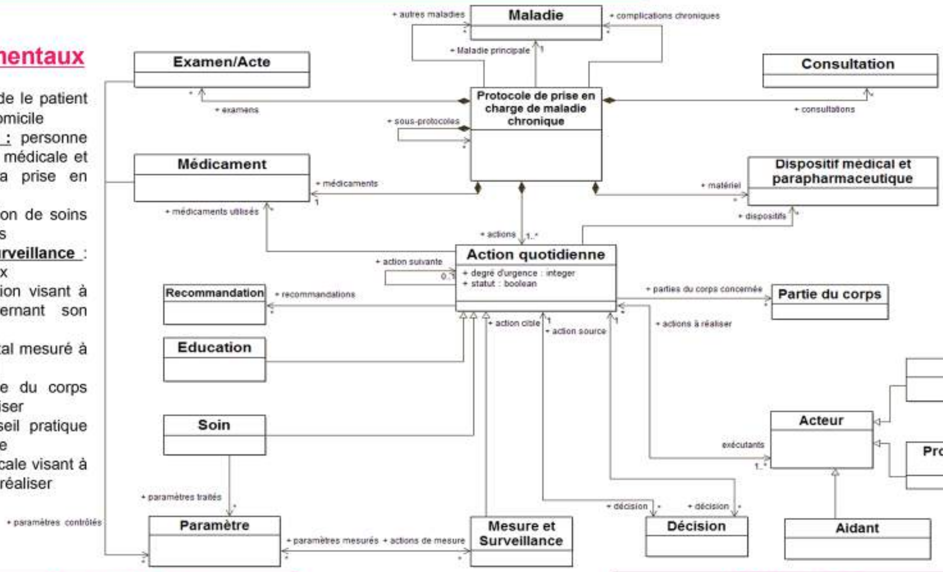
Verrous scientifiques

- Comment construire des protocoles de soins **personnalisés** et **faciles à comprendre** par le patient ?
- Comment **intégrer** et filtrer les imprévus dans un protocole de soins ?
- Comment **automatiser** le protocole de soins et l'**adapter** en fonction des imprévus pour permettre son **auto-enrichissement** ?

Un méta-modèle et un langage orientés patient

Concepts fondamentaux

- **Aidant :** personne qui aide le patient dans sa prise en charge à domicile
- **Professionnel de santé :** personne qui appartient à une équipe médicale et qui est impliquée dans la prise en charge à domicile
- **Action de soin :** réalisation de soins préventifs, curatifs et palliatifs
- **Action de mesure et surveillance :** contrôle de paramètres vitaux
- **Action d'éducation :** action visant à éduquer le patient concernant son protocole
- **Paramètre :** paramètre vital mesuré à l'aide d'un dispositif médical
- **Partie du corps :** partie du corps concernée par l'action à réaliser
- **Recommandation :** conseil pratique et information supplémentaire
- **Décision : décision** médicale visant à choisir la prochaine action à réaliser



Langage graphique

Perspectives

- Filtrage et adaptation du protocole de soins : utilisation d'ontologies pour représenter les imprévus
- Amélioration de l'IHM dédiée au patient en utilisant des dispositifs de reconnaissance de la parole
- Implémentation de la solution logicielle sur un dispositif mobile en collaboration avec Arcan Systems

Publications

- Derradji A. (2014). Construction de workflow auto-adaptatif pour le suivi des maladies. Forum Jeunes Chercheurs du 32e congrès INFORSID. Lyon, France.
- Derradji A. (2015). Construction de protocoles auto-adaptatifs de soins pour le suivi des maladies. À paraître dans la revue Ingénierie des Systèmes d'Information,

Vers un outil auteur pour des EIAH destinés à l'apprentissage de méthodes de résolution de problèmes

Awa Diattara [a, b], Equipe MeTAH (LIG) [a] & TWEAK (LIRIS) [b]
Vanda Luengo [a], Nathalie Guin [b]

CONTEXTE

❖ **LE PROJET AMBRE**

Conception d'EIAH (Environnements Informatiques pour l'Apprentissage Humain) destinés à favoriser l'apprentissage de méthodes de résolution de problèmes.

❖ **LES EIAH AMBRE**

- Proposent un processus d'apprentissage inspiré du RàPC : cycle AMBRE
- Offrent à l'apprenant des fonctionnalités d'aide, de diagnostic et d'explications sur ses erreurs

OBJECTIF

Permettre à un expert non informaticien (un enseignant par exemple) d'expliciter les connaissances nécessaires à la conception d'un EIAH AMBRE

QUESTIONS DE RECHERCHE

1. Acquisition de connaissances suffisantes pour faire du raisonnement
   - ❑ Proposition de méta-modèles de connaissances
   - ❑ Conception des interfaces permettant d'instancier ces méta-modèles
2. Comment concilier l'acquisition des connaissances avec la conception de l'interface de l'EIAH ?

TRAVAUX ENGAGÉS/ RÉSULTATS

- Formalisation des connaissances à acquérir : proposition de méta-modèles à l'aide de schémas XML (un article publié à la conférence TICE)
- Conception de l'interface de l'outil auteur : acquisition des connaissances sur la méthode à enseigner

PERSPECTIVES

- Validation des méta-modèles (Tests sur d'autres domaines)
- Tests de l'utilisabilité de l'outil auteur auprès d'enseignants (expérimentations)

POINTS CLES

Assistance à l'utilisateur

- ❑ Flexibilité : ordre et format de définition
- ❑ Généralisation à partir d'exemples

METAH "Modèles et Technologies pour l'Apprentissage Humain"

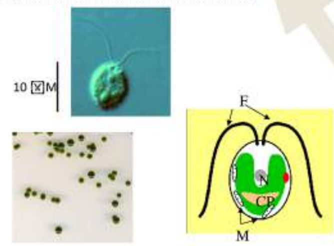
Design of a Bayesian machine

Marvin Faix - CNRS - marvin.faix@inria.fr
supervisor : Emmanuel Mazer
co supervisor : Laurent Fesquet

"Perception, reconnaissance et intégration pour la modélisation d'activité" PRIMA

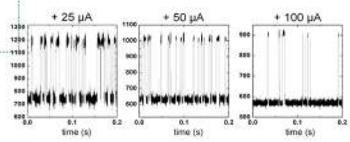
BAMBI – Bottom-up Approaches to Machines dedicated to Bayesian Inference

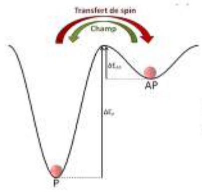
«bio study chlamydomonas»

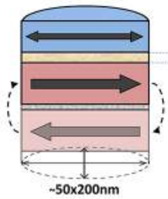


MTJ NANO COMPONENTS :

- intrinsically stochastic nano component
- tunnable
- coding in time







Soft evidence

$P(MDL)$: def

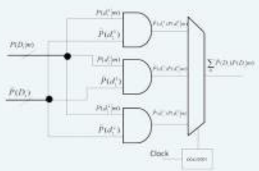
$\tilde{P}(D_i)$: soft evidence input

$$P(M|\prod_i \tilde{P}(D_i)) = \frac{1}{Z(D)} \sum_{D_1} \tilde{P}(D_1) \cdots \sum_{Dk} \tilde{P}(D_k) \sum_{L} P(MDL)$$

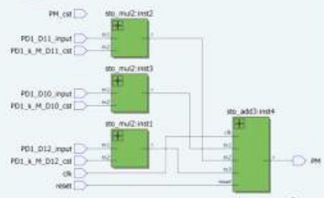
Bayesian model description

$$Pr \begin{cases} Ds \begin{cases} Sp(\pi) \begin{cases} Va : D_1, D_2, M \\ Dc : \{P(MD1D2) = P(M)P(D_1|M)P(D_2|M) \\ Fo : \begin{cases} P(M) \text{Uniform} \\ P(D1|M) \text{Histograms} \\ P(D2|M) \text{Histograms} \end{cases} \end{cases} \\ Identification \end{cases} \\ Qu : P(M) \end{cases}$$

high level circuit simulation



mapping on FPGA



Stochastic computation

BITSTREAM REPRESENTATION

> *Gaines, Stochastic computing systems, Advances in Information Systems Science vol 2 37-172 1969.*

$$010001101100 = \frac{5}{12}$$

--> strong lengths

- low power
- fault tolerent :

classic representation:
01001010 = 74

If bit flip on MSB:
11001010 = 202

8 bit stream value representation:
01001010 = 3/8

If bit flip:
11001010 = 4/8

=> Each bit has the same weight !

--> weaknesses

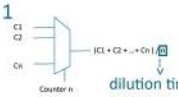
- low accuracy
- long time computation

STOCHASTIC ARITHMETIC

X

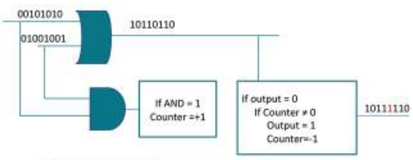
\+

method 1



dilution time probem

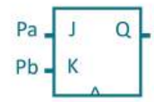
method 2



$R = 0$	output	R
00	0	0
01	1	0
10	1	0
11	1	1

$R \neq 0$	output	R
00	1	$R-1$
01	1	R
10	1	R
11	1	$R+1$

÷



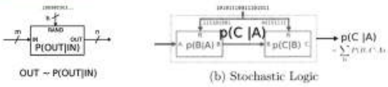
$$Pc = \frac{Pa}{Pa + Pb}$$

normalization

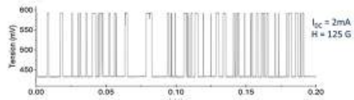
OTHER APPROACHES

sampling method

> *Jonas E., Stochastic Architectures for Probabilistic Computation, PhD Dissertation 2014.*

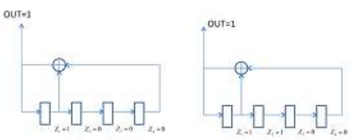


telegraphic signal



LFSR pseudo noise acquisition

LFSR SYNCHRONIZATION



$Z_1(t) = Out(t-1)$
$Z_2(t) = Z_1(t-1)$
$Z_3(t) = Z_2(t-1)$
$Z_4(t) = Z_3(t-1)$

$Out(t) = Z_1(t) \oplus Z_4(t)$
$Z_1(0), Z_2(0), Z(0)_3, Z_4(0) \neq 0$

BAYESIAN MODEL

$$P(S_T|\tilde{P}_T(O_T)) = P(S_T|\tilde{P}_1(0_1)\ldots\tilde{P}_T(0_T)) = \frac{1}{Z}\Big(\sum_{O_T}(\tilde{P}_T(O_T)P(O_T|S_T)\Big)\sum_{S_{T-1}}(P'(S_{T-1})P(S_T|S_{T-1}))$$

with

$$P'(S_{t-1}) = P(S_{t-1}|\tilde{P}^*_{t-1}(O_{t-1}))$$

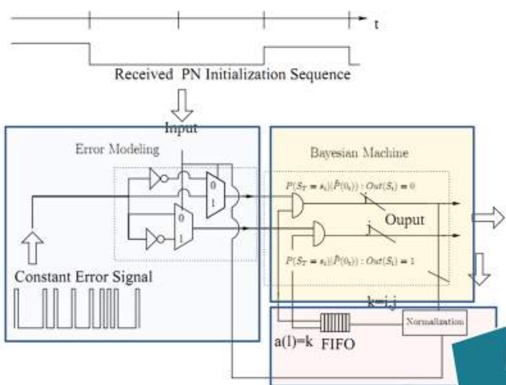
and

$$P(S_T = s|\tilde{P}_T(O_T)) = \frac{1}{Z}P'(a(s))\tilde{P}_T(Out(s))$$

NOISY TRANSMISSION

Register	state	output	received
[1, 0, 0, 0]	1	1	1
[1, 1, 0, 0]	3	1	1
.[1, 1, 1, 0]	7	1	1
[1, 1, 1, 1]	15	1	0
[0, 1, 1, 1]	14	0	0
[1, 0, 1, 1]	13	1	1
[0, 1, 0, 1]	10	0	0
[1, 0, 1, 0]	5	1	1
[1, 1, 0, 1]	11	1	1
[0, 1, 1, 0]	6	0	1
[0, 0, 1, 1]	12	0	0
[1, 0, 0, 1]	9	1	0
[0, 1, 0, 0]	2	0	0
[0, 0, 1, 0]	4	0	0
[0, 0, 0, 1]	8	0	1
[1, 0, 0, 0]	1	1	0

STOCHASTIC CIRCUIT



Towards Matching Improvement between Tasks and Workers in Spatial Crowdsourcing Systems

André Sales Fonteles

Advisors: Jérôme Gensel and Sylvain Bouveret

Context

Crowdsourcing systems are platforms that enable a system or a user to publish tasks that other users, called workers, are intended to accomplish. We focus on Spatial Crowdsourcing Systems such as the following scenario:

Matching Points of View

System

Try to assign at least one worker to each task in the system, respecting some constraints.

Drawback: Workers may be assigned to tasks they dislike, rendering bad or no contribution

Worker

Try to assign or recommend tasks to a worker that best match his preferences, ignoring a global optimality of the system.

Drawback: Tasks that tend to be not attractive for workers may suffer starvation.

Task

Try to assign or recommend tasks to a worker that best matches his skills/ abilities, ignoring a global optimality of the system.

Drawback: Tasks that require unusual skills from a worker may suffer starvation.

Objectives

Help workers to find spatiotemporal tasks, and/or a sequence of them, having as goals:

- To improve the overall contribution of a worker.
- To increase the quality of the service provided by workers.

Proposal

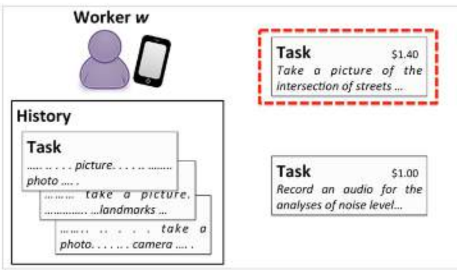
Single Task Recommendation

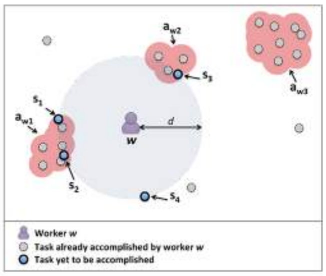
(Fonteles et al., 2014, MobiGIS) (Fonteles et al., 2014, SAGEO)

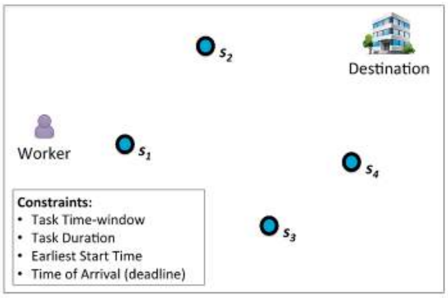
Estimate the utility *u(s,w)* of a task *s* for a worker *w*.

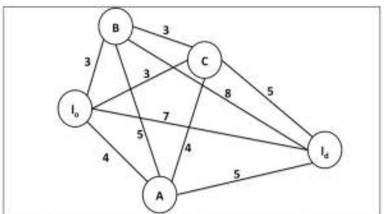
- *Model Proposed*
  - *General preferences*
  - *Skills*
  - *Spatiotemporal Preferences*
  - *Reward*

Rank tasks according to their utility for recommendation









Task	Time window ($[\tau_1,\tau_2]$)	Duration (δ)	Utility (u)
A	[1:00 pm, 1:50 pm]	5 minutes	0.5
B	[1:00 pm, 2:10 pm]	15 minutes	0.5
C	[1:30 pm, 2:40 pm]	10 minutes	0.6

Task Sequence Recommendation

(Fonteles et al., 2015, W2GIS)

Find and recommend a valid optimal sequence that maximizes utility.

- Problem formalized
- Problem proved to be NP-hard
- Exact algorithm proposed and experimented
- Approx. algorithms proposed and experimented

List of Publications

Fonteles, A. S., Bouveret, S., & Gensel, J. (2014, November). Towards matching improvement between spatio-temporal tasks and workers in mobile crowdsourcing market systems. In Proceedings of the Third ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems (pp. 43-50). ACM.

Fonteles, A. S., Bouveret, S., & Gensel, J. (2014, November). Améliorer l'appariement entre tâches et exécutants dans les Systèmes de Marché Participatifs Mobiles. In Proceedings of the International Conferece SAGEO. Grenoble, France.

Fonteles, A. S., Bouveret, S., & Gensel, J. (**to appear**). Opportunistic trajectory recommendation for task accomplishment in crowdsourcing systems. in International Symposium on Web and Wireless Geographical Information Systems. Grenoble, France, May 2015.

Fonteles, A. S., Bouveret, S., & Gensel, J. (**submitted**). Approximation Algorithms for Task Recommendation in Crowdsensing Systems. in International Conference on Distributed Computing in Sensor Systems. Fortaleza, Brazil, June 2015.

Fonteles, A. S., Bouveret, S., & Gensel, J. (**submitted**). Recommandation opportuniste de trajectoires pour l'accomplissement de tâches dans les systèmes crowdsourcing. in Proceedings of the Conference INFORSID. Biarritz, France, May 2015.

Future Work

- Try to also take into account the system's point of view (global optimum).
- Propose more approximation algorithms for the Task Sequence.
- Investigate the possibility to adapt TSP algorithms to the Task Sequence.

STEAMER

"Spatio-temporal information, adaptability, multimedia and knowledge representation"

Efficient algorithms to mine per-item patterns in user-generated data

Martin Kirchgessner, advised by Sihem Amer-Yahia & Vincent Leroy

Preliminary works with Alexandre Termier & Marie-Christine Rousset

We adapt pattern mining to user-generated data, at Web scale.

- Many systems record per-user (or per-action) lists of items: retail tickets, playlists, browser logs...
- Patterns (here, sets of artists) allow an analyst to: GOALS
  - summarize the data by highlighting co-occurrences
  - link each pattern back to the concerned users, for further analysis

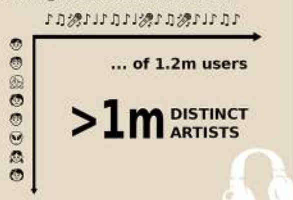
Frequent Itemset Mining can find that
- 294,720 people listen to Radiohead
- 61,487 to {The Beatles, The Rolling Stones}
- 73,666 to {Coldplay, Muse} .... and millions of such sets !

Lesser-known artists are also part of interesting patterns ! CHALLENGES

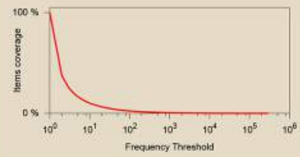
FIM can't find
- The 19,882 people who all listen to {Franz Ferdinand, Radiohead, Interpol} ?
- 1,995 hipsters listening to {Vitalic, Hot Chip} ?
- 1,994 (metal) fans of {Children of Bodom, In Flames, Arch Enemy, Eternal Tears of Sorrow}
- 200 (likely French) listeners of {Tryo, Les Wriggles, Fabulous Trobadors}
- 10 eclectic users who know {Sido, B-Tight, Tony D, G-Hot, Fler, Sido, Alpa Gun} ?

EXAMPLE : LastFM

Listing the 50 favorite artists...



... but most artists are far from frequent



We propose

TopPI

The first algorithm and implementation to find Per-Item patterns

READABILITY ASSUMPTIONS

1. It's OK to study 2 or 3 dozens of patterns
2. The analyst usually jumps in the results via a key item

"What's related to The Beatles ?"

A NEW PROBLEM STATEMENT

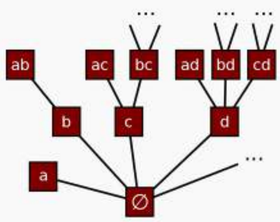
Given transactions over a set of items I and an integer k, return, $\forall i \in I$, the k most frequent closed itemsets containing i.

↳ We get, for each artist a listened to at least 2 times, k sets of artists including a

MINING, ALL AT ONCE

WHY Mining {a,b,c} is much more profitable if it's a top-pattern for a, b and c

HOW
- rely on existing mining methods [1][2]
- explore top-per-item solutions first
- prune in a parallel-friendly manner



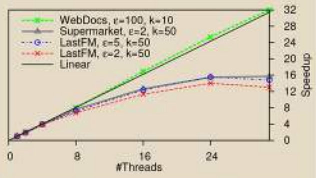
BIG DATA SYSTEMS

Modern systems are heavily parallel. Our algorithms leverage such architectures, to analyze datasets at all scales.

On multi-core

TopPI's solution space is a tree.

By assigning a branch to each available core [2], we achieve good speedups :

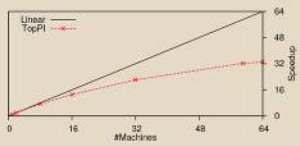


Examples presented here are extracted from LastFM's top-50-per-item, computed in ~15 minutes on the author's laptop.

On clusters

We partition items among machines.

Completeness is ensured by splitting the mining in 2 phases, during which each machine works independantly.



FUTURE WORK: MINING 1.4 BILLION RECEIPTS

Will the data partitoner handle so many lines ?

RANKING BY CORRELATION FUTURE WORK

Very frequent items tend to add noise to patterns.

Statistical correlation appears to be better than support at sorting the best per-item patterns [3]

What if we change the ranking: find, for item a, the k itemsets mostly correlated to a?

CLOSURES ARE NOT ENOUGH

TopPI only gives *closed* itemsets [1] :

123 {a,b,c}

~~123 {a,b}~~

~~123 {a}~~

But each itemset may not fit directly as a recommended set :

63030 {Radiohead, Bloc Party}
62962 {Radiohead, Beck
61197 {Radiohead, Modest Mouse}

[1] T. Uno, M. Kiyomi, and H. Arimura. *Lcm ver. 2: Efficient mining algorithms for frequent/closed/maximal itemsets.* FIMI, 2004

[2] B. Négrevergne, A. Termier, M.-C. Rousset, and J.-F. Méhaut. *Paraminer: a generic pattern mining algorithm for multi-core architectures.* DMKD, 2013

[3] G. Liu, M. Feng, Y. Wang, L. Wong, S-K. Ng, Tzia Liang Mah, E.J.D. Lee *Towards exploratory hypothesis testing and analysis* ICDE, 2011

SLIDE "ScaLable Information Discovery and Exploitation"

Laboratoire d'Informatique de Grenoble 2015 LIG

Every pattern in this poster is based on real data
PhD funded by Projet Datalyse (datalyse.fr) / Investissements d'Avenir, développement de l'économie numérique

Spatio-Temporal Concept Drift Exploration

Sofia Kleisarchaki
Advisors: Sihem Amer-Yahia, Ahlame Douzal-Chouakria, Vassilis Christophides

Context

"Universe of Data": data form concepts evolving in spacetime

Goals

1. Observe the universe to reveal *concept drifts*
2. Explore the observable space to query time and space dynamics of drifts

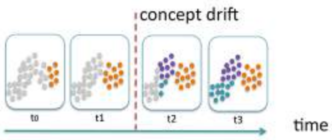
Challenges

1. Evolutionary data from varying distributions
   - Need for dynamic learning of concept drift parameters
2. Dynamic drifts of *varying arrival rates*
   - Need to explore multiple time and space granularities
3. Data of *high volume*
   - Need for scalable structures and efficient query exploration algorithms

Problem

Given a dataset D and a period p, detect all concept drifts of D at any granularity g of p.

Concepts of D in colorful clusters:

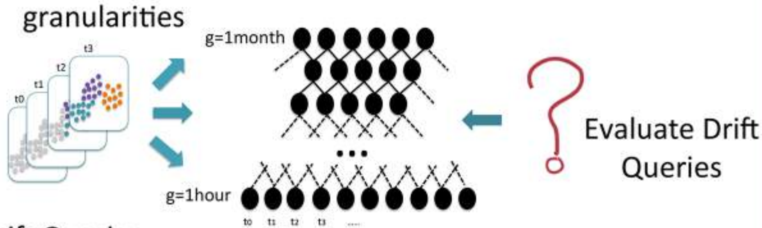


Contributions

1. Scalable drift index maintaining concepts in multiple granularities
2. Formalize drift queries in order to provide flexibility in tracking and analyzing drifts in evolving data
3. Efficient evaluation of drift queries

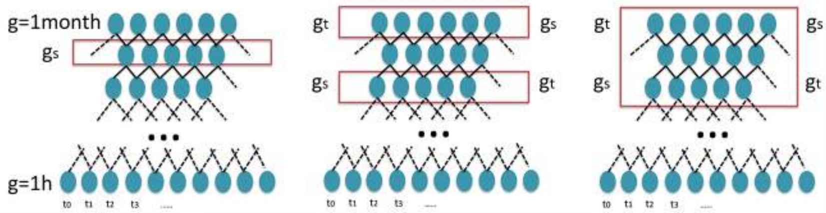
Our Approach

- Drift Index: Directed graph G=(V, E)
  - V: Set of clustering nodes at different time intervals of p
  - E: Set of edges between nodes of consecutive granularities



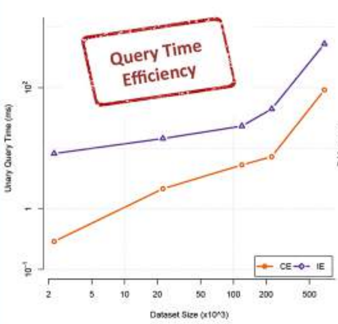
- Drift Queries

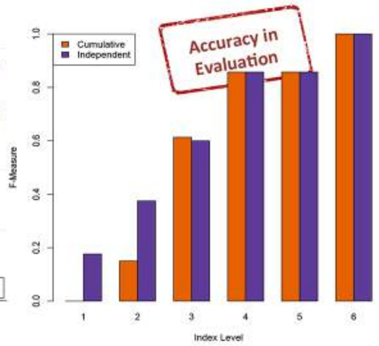
Unary	***Binary***	***Range***
drifts at g_s	drifts at g_t included in g_s	drifts at $[g_s, g_t]$



Experiments

Datasets: Usenet (below), KDD-Cup'99, Twitter





Future Work

Spatio-Temporal Concept Drifts

- Drifts occur at several geographic levels (e.g., Europe, France) and time granularities (e.g., day, month)
- Provide a flexible approach for querying spatio-temporal concept drifts

Contact Info

"ScaLable Information Discovery and Exploitation" SLIDE

Laboratoire d'Informatique de Grenoble 2015 LIG

[1]: S.Kleisarchaki, S. Amer-Yahia, A. Douzal-Choukria, V. Christophides. Online Detection of Topic Change in Social Posts. Big'14
[2]: A Flexible Approach for Querying Topic Drifts, "Under preparation".

Quentin Labourey Directeurs : Olivier Aycard et Denis Pellerin

Développement d'un robot attentionné pour la surveillance de personnes en situation de fragilité

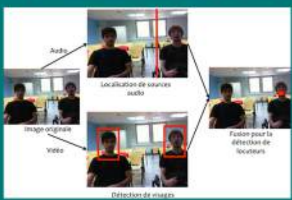
Problématique :

Surveillance par des capteurs dans l'infrastructure mal acceptée (caméras, micros, etc…)

Robot compagnon attentif capable de :
- Percevoir son environnement
- Analyser des situations complexes
- Focaliser son attention
- Naviguer dans un environnement dynamique

Qbo :
- Microphones
  - Kinect
  - Joues
- Camera :
  - Yeux (RGB)
  - Kinect (RGBD)
- Actionneurs :
  - Tête
  - Base

Détection et suivi de locuteurs successifs par la fusion audiovisuelle [1] :



Audio : Localisation de sources audio grâce au « Time-Difference of Arrival » suivi d'une cross-correlation

Vision : Détection de visage, apprentissage sur Viola et Jones et détection à l'aide de filtres bayésiens naïfs

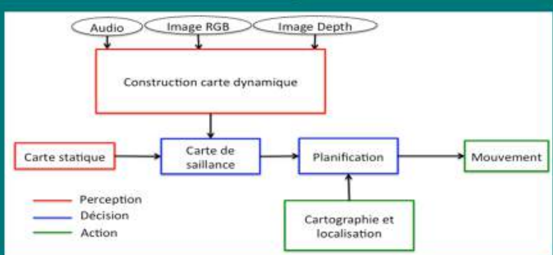
Fusion : Détection et suivi du locuteur actuel à l'aide d'un filtre bayésien classique

Développement d'un système de surveillance humaine basé sur la fusion audiovisuelle et la navigation :

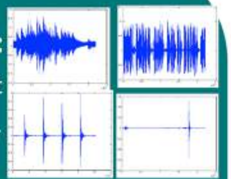
Le but ici est de surveiller et « tagger » les humains présents dans une pièce en naviguant de manière intelligente.

2 cartes de navigation :
- Une carte statique : les endroits à visiter par défaut en patrouille
- Une carte dynamique : modifier l'itinéraire en temps réel selon les observations capteurs

Comment choisir la stratégie de navigation pour visiter tous les points d'intérêt de manière intelligente ?

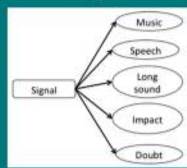


Classification de sons en environnement intérieur [2] :



Les sons entendus en environnement intérieur (e.g. appartement, maison) peuvent donner des informations sur le contenu d'une scène.

Problème : Contenu très riche et complexe, sons « proches » ayant des sens très différents. Il faut donc choisir une taxonomie adaptée au problème multimodal dans lequel on s'inscrit : il faut pouvoir laisser la place au doute



Extraction de features « classiques » de la littérature : Mel-Frequency cepstrum coefficients (MFCC), Flux spectral, Taux de trames à basse énergie, etc…

Classification adaptée au cadre multimodal : Belief K-NN : construction d'une fonction de croyance pour chaque échantillon s à classifier parmi les classes (C_1,..,C_N) selon ses K plus proches voisins.

Pour chaque voisin de classe C_q, une masse de croyance est construite :

$$\begin{cases} m^s(C_q) = \alpha f(d) \\ m^s(C_i) = 0 \quad \text{si } i \neq q \\ m^s(C) = 1 - \alpha f(d) \end{cases}$$

Puis les masses de croyance sont aggrégées pour former une masse finale, dont le maximum est sélectionné pour choisir la classe

Test sur une base de données tirée de la vie réelle.

Class	Nb of training extracts	Nb of test extracts
Speech	11	16
Music	11	19
Impact	5	11
Long Sounds	5	12
Total	32	58

Résultats sur 1000 cross-validations

	Speech	Music	Impact	Long	Doubt	Total
KNN	1.4	1.3	0.7	1.1	4.1	8.6
SVM	0.6	2.5	2	1.1		6.3

Nombre d'échantillons mal classés

Références :

[1] Labourey Q., Aycard O., Pellerin D., Rombaut M., Audiovisual data fusion for successive speaker tracking, International Conference on Computer Vision Theory and Applications (VISAPP 2014), Lisbon, Portugal, January 2014

[2] Labourey Q., Pellerin D., Rombaut M., Aycard O., Garbay C., Indoor sound classification based on belief functions, European Signal Processing Conference (EUSIPCO 2015) (en cours de soumission)

AMA

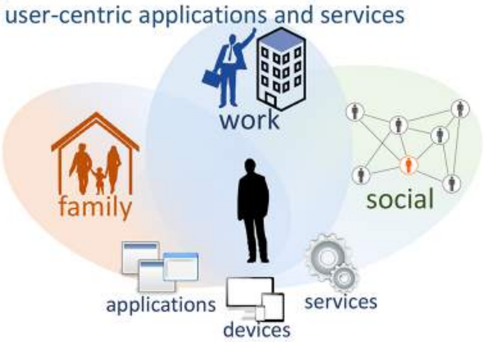
"Analyse de données, Modélisation et Apprentissage automatique"

Contextualization for the Personal Cloud

My Personal Cloud: a place where I interact in safe with my digital life

Goal: To facilitate the management, the execution and the interaction of the user with his/her various digital lives (professional, personal, cultural, social, etc.)

How: Context-awareness to provide personalized user experience with user-centric applications and services



Context-Awareness

- A service can be executed in various forms ↔ various Quality of Experience (QoE) levels
- Each user expects his own QoE when using a service
- To ensure a QoE level, the service has its specific resource requirements

Resource(availabilities, specificities) + User's profile(preferences, expertness)

= Execution Context

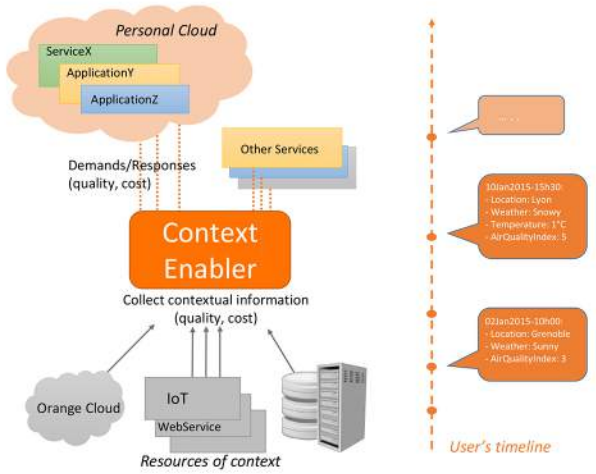
- Service deployment optimization, personalization:
  Ex: location-based cache for data-sharing services
- Adaptation, reconfiguration of the execution:
  Ex: video encoding adaptation based on network quality

Feeding the Personal Cloud with a quality and cost-aware Context-Enabler

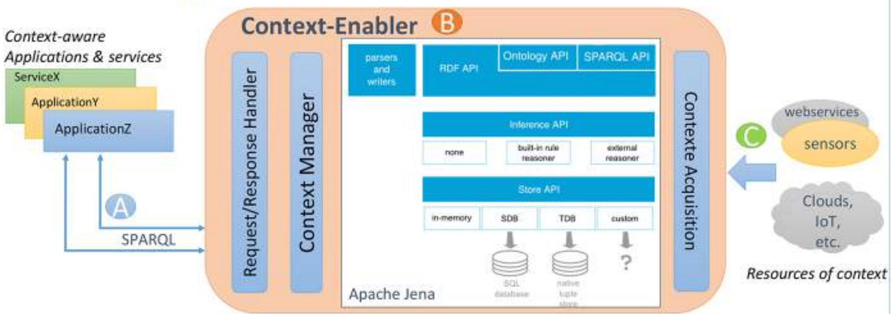
(A) Context-aware applications express their demands of context (with quality and cost constraints) to the Context-Enabler.

(B) The Context-Enabler uses semantic description, quality and cost information to classify and to make choices of context resources providing semantically equivalent information but different in quality and cost.

(C) The Context-Enabler is able to provide a continuous response to the client by adapting dynamically its context provisioning chain regarding the changes happening during its execution.



Prototype



Context representation: Ontology OWL/RDF
RDF framework: Apache Jena
Rules of reasoning: SWRL

Query language: SPARQL, C-SPARQL
Communication/Messaging : RabbitMQ

Research perspectives

- Definition and implementation of the Quality and Cost of context aspects
- Representation and manipulation of semantically equivalent resources
- Language to query and to handle complex-context information
- Techniques and mechanisms that allow the dynamic adaptation of the Context-Enabler

PhD Student
Anh Dung LE
anh-dung.le@imag.fr

LIG – SIGMA
Claudia RONCANCIO, Cyril LABBE
claudia.roncancio@imag.fr cyril.labbe@imag.fr

Orange Labs
François-Gaël OTTOGALLI
francoisgael.ottogalli@orange.com

orange

SIGMA

"Systèmes d'Information - inGénierie et Modélisation Adaptables"

Contiguity and Locality in Backfilling Scheduling

Fernando Mendonca, supervised by Denis Trystram and Frederic Wagner

Introduction

High Performance Computing systems are evolving from large scale multi-cores to extreme scale heterogeneous many-cores (where hundreds of cores are integrated within the same chip). The number of cores will drastically increase but the I/O and interconnection devices are evolving much slowly while the memory hierarchy will be even deeper than today.
The most popular scheduling mechanism in supercomputing centers is First Come First Served (FCFS). It consists in executing the jobs in their order of arrival and to allocate them in the first available time slot. FCFS is used with local improvements aiming at filling idle times with smaller jobs in the queue (Backfilling). Although Backfilling does not optimize any sophisticated function, it provides good results due to high resource occupation, it is simple to implement and it guarantees that there is no starvation.

Related works

Backfilling algorithms are the most commonly used scheduling algorithms in batch schedulers. Conservative backfilling is the most commonly version mentioned in the literature. In this version, backfilling is done only if it does not delay *any* of the previous jobs in the queue. Users provide each submitted job with a running time limit and a fixed number of processors.
Pascual et al. proposed *topology-aware* backfilling techniques in. They introduced modifications to processors allocation algorithms in order to improve job locality in k-ary n-trees platforms topologies. They show an increase in waiting times for strict locality enforcing.
Consider the example of a fat tree topology as displayed in Figure 1. Two different cluster configurations can be mapped on the tree hierarchy. We can choose to define clusters at the lowest hierarchical level to obtain four clusters $(C_1, \ldots, C_4)$ of size two. A second choice is possible by dividing one level above into two clusters (C'_1, C'_2) of size four.

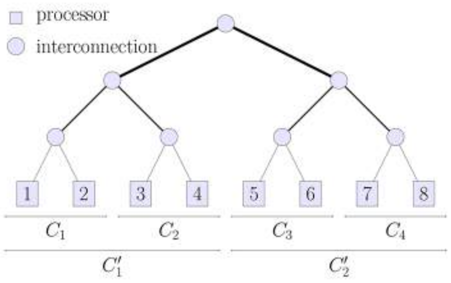


Fig. 1: Example of a fat tree interconnection

Backfilling

Conservative backfilling uses two data structures: a list of queued jobs and the times at which they are guaranteed to start execution. The other is a profile storing the status of processors for each of the future periods of time. These time slices are potentially created/destroyed on job insertion.

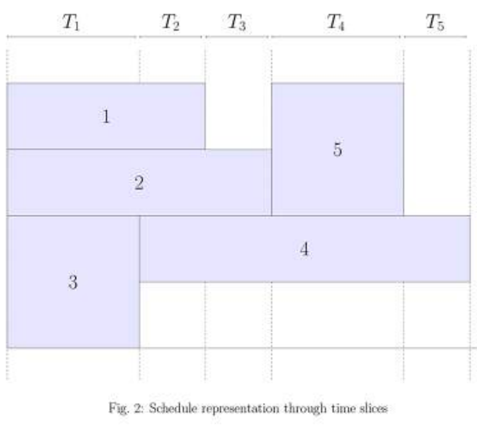


Fig. 2: Schedule representation through time slices

When scheduling a job, the backfilling algorithm starts by finding the first contiguous set of time slices large enough for both the number or required processors and the execution time.
One of the final steps of the backfilling algorithm (as shown in Algorithm 1 line 15) is to reduce the set P of available processors. This operation is called the *allocation* step.
All of our algorithms are based on the specialization of this step in order to select specific processors. We implement the following algorithms: **basic, best effort contiguous, forced contiguous, best effort local, forced local and forced local contiguous**.

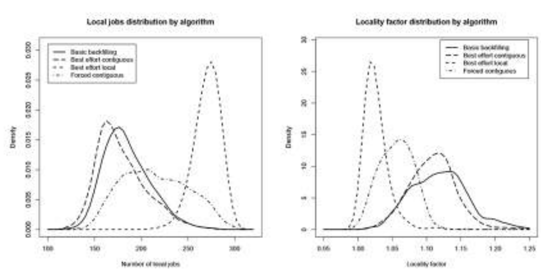
Algorithm 1: Backfilling

Data: job j, processing time p_j, required processors number $size_j$, list of time slices S
Result: list P of available processors for the job, starting time t

1 **foreach** *time slice* $S[i]$ **do**
2   **if** $|S[i]| < size_j$ **then**
3   next;
4   **end**
5   $P \leftarrow S[i]$;
6   $time \leftarrow 0$;
7   **foreach** *time slice index j starting from i* **do**
8     $P = P \cap S[j]$;
9     $time \leftarrow time + duration(S[j])$;
10   **if** $time \geq p_j$ **then**
11   break;
12   **end**
13   **end**
14   **if** $|P| \geq size_j$ **then**
15     $P \leftarrow size_j$ processors of P;
16     $t \leftarrow starting\_time(S[i])$;
17   **return** (P, t)
18   **end**
19 **end**

Experimental results

We study locality which is the main target of our work. In the first Figure below the probability density function of the number of local jobs is plotted. We can see that *best effort contiguous* achieves worse results than *basic*. *Forced contiguous* outperforms *basic* with however more variability in the performance. Finally *best effort local* achieves the best results by scheduling jobs locally 90% of the time. In order to get a better view of the achieved locality of jobs we propose to consider a new metric, the *locality ratio*.
The Figure below displays the density function of locality ratios for all algorithms. *Forced local* is not displayed in this figure since all its jobs are local. Figure 3 shows that with respect to the locality ratio all our algorithms proceed as expected. *Best effort local* gives the best results, followed by *forced contiguous*. *Basic* and *best effort contiguous* show similar performances with a slight advantage for *best effort contiguous*.



Acknowledgment

This paper was partly funded by CAPES Foundation, Ministry of Education, Brazil. It was also given support from the ANR project MOEBUS.

MOAIS
"Multi-programmation et Ordonnancement sur ressources distribuées pour les Applications Interactives de Simulation"

Human-Robot Motion: an Attention-Based Navigation Approach

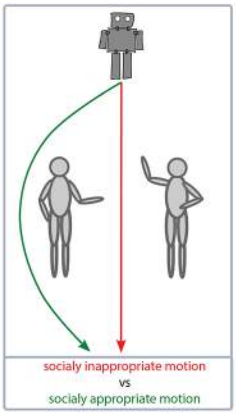
Rémi Paulin

Problem: robot motion amongst humans

Sharing the environment with humans

1980s- 2005: Safety & optimality

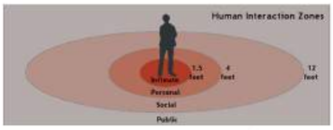
Robots can efficiently avoid obstacles, but it is not enough.



People are not pieces of furnitures; the motion in red is shorter, but not socially appropriate.

State-of-the-art: social spaces

Depending on the activity humans are involved in, the robot should avoid certain regions of space where its presence would cause discomfort to the humans.

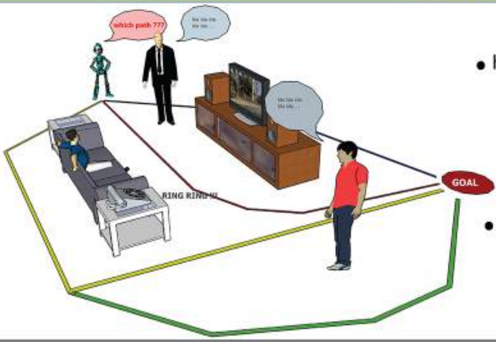


personal space[1]

interaction space[2]

activity space[3]

Towards more complex interactions



- humans are involved in complex interactions
  - multiple interactions and activities (watching TV, conversation with two individuals)
  - non-intended distracting objects (phone, robot)
- complex tasks for the robot
  - how to get everyone's attention?
  - how to deliver a message to one person with the least amount of distraction to others?
  - how to accomplish its task while minimising the amount of distraction to everone?

Attention: a new tool for navigation

Depending on an individual's intention (i.e. intended activity) and the distractions in their environment, they will distribute *attentional ressources* to both intented interaction/activities and unintended sources of distraction.

attention model[4]

An *attention vector* encapsulates the individual's intention and the influence of their environment. It is defined by:

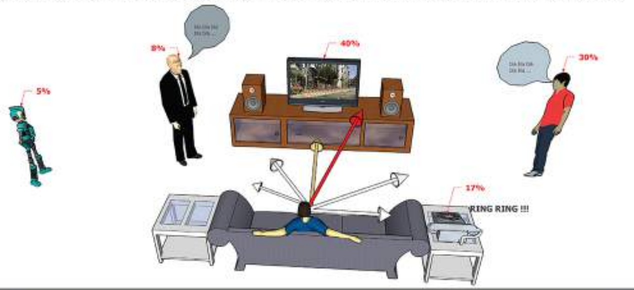
Attention = f(Intention, Distractions)

This *attention vector* contains information on the general direction of the individual's *attention focus*. The *attentional resources* are then allocated to the distractions of the scene, relative to the main direction of attention.

		8%	30%	5%	40%	17%
	45%		2%	35%	8%	10%
	60%	3%		20%	2%	15%

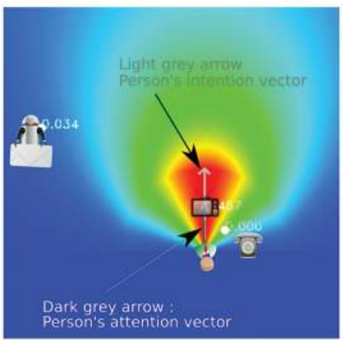
The *attention matrix* shows the amount of *attentional resources* given by each individual (lines) to each distraction (columns) in the scene.

A case study: the values in red show the percentage of *attentional resources* given by the individual sitting on the sofa to each distraction in the scene.



First results

Our navigation problem can be formulated in terms of optimisation of the *attentional resources* received by the robot along its trajectory.



The *attentional resources* allocated by the individual to the robot depend on its position, thus defining an attention field. Warm colors indicate regions of greater distraction and should generally by avoided by the robot - unless it is specifically seeking attention.

"Perception, reconnaissance et intégration pour la modélisation d'activité"

PRIMA

Laboratoire d'Informatique de Grenoble

2015

[1] Hall, E. T., & Hall, E. T. (1969). The hidden dimension (Vol. 1990). New York: Anchor Books.
[2] Lindner, F., & Eschenbach, C. (2011). Towards a formalization of social spaces for socially aware robots. In Spatial Information Theory (pp. 283-303). Springer Berlin Heidelberg.
[3] Bakker, S., Van Den Hoven, E., & Eggen, B. (2012). Knowing by ear: leveraging human attention abilities in interaction design. Journal on Multimodal User Interfaces, 5(3-4), 197-209.
[4] J. Maisonnasse (2007), Estimation des relations attentionnelles dans un environnement intelligent. PhD Thesis, University Joseph Fourier, Grenoble (FR).

Supervision: Thierry Fraichard et Patrick Reignier

From efficient Byzantine Fault Tolerance to efficient Fault Tolerance

Lucas Perronne

Introduction

- Cloud Computing allows the design of reliable and efficient applications on distributed resources
- In such settings, State Machine Replication is a suitable way to provide Fault Tolerance
- We show interest in Byzantine Fault Tolerance, designed to tolerate arbitrary and malicious behaviors

Motivation

- Byzantine Fault Tolerance protocols provide dependability guarantees under specific assumptions
- Protocols focus on optimizing performance while holding these guarantees

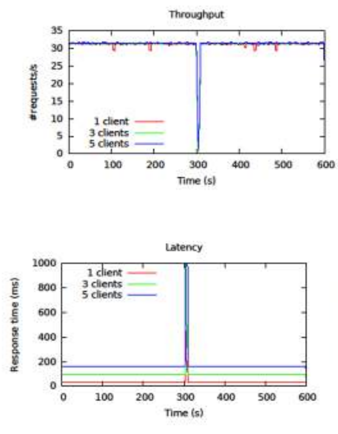
Objective

Design a BFT protocol that adapts its communication pattern to the detected faults

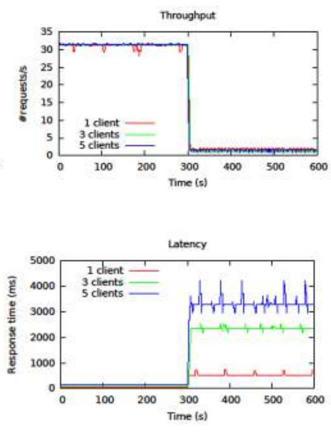
Target Contributions

- Classify faulty behaviors handled by BFT into subsets of increasing complexity
- Propose a sub-protocol for each subset while maintaining the BFT dependability guarantees and improving performance

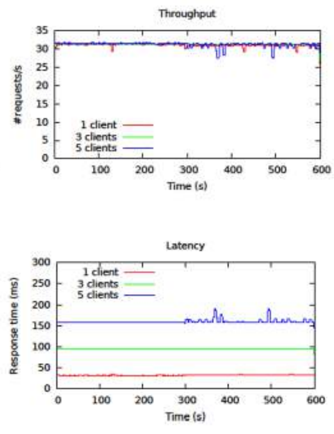
Crash



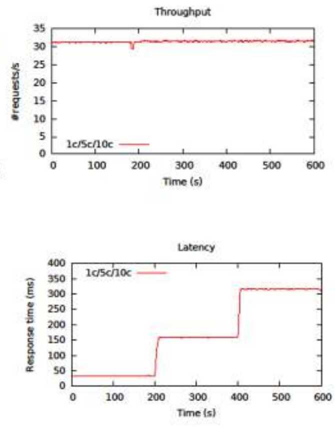
Delay



Flood



1c/5c/10c



Acknowledgements

- AMADEOS (Architecture for Multi-criticality Agile Dependable Evolutionary Open System-of-Systems) is a collaborative project funded under the European Commission's FP7 (FP7-ICT-2013-610535). Project duration 2013-2016.
- Part of the experiments were conducted on the Grid'5000 experimental testbed, developed under the INRIA ALADDIN development action with support from CNRS, RENATER and several Universities, as well as other funding bodies.

References

[1] M. Castro and B. Liskov. Practical Byzantine fault tolerance. In Proceedings of the 3rd Symposium on Operating Systems Design and Implementation, Feb. 1999.

[2] R. Kotla, L. Alvisi, M. Dahlin, A. Clement, and E. Wong. Zyzzyva: speculative Byzantine fault tolerance. Technical Report UTCS-TR-07-40, University of Texas at Austin, Austin, TX, USA.

[3] Algirdas Avizienis, Jean-Claude Laprie, Brian Randell, and Carl E. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Trans. Dependable Sec. Comput.*, 1(1) :11–33, 2004.

Laboratoire d'Informatique de Grenoble

2015 LIG

ERODS

"Efficient and RObust Distributed Systems"

Validation Conjointe en UML et B de Modèles de Sécurité en SI

Par: Amira RADHOUANI

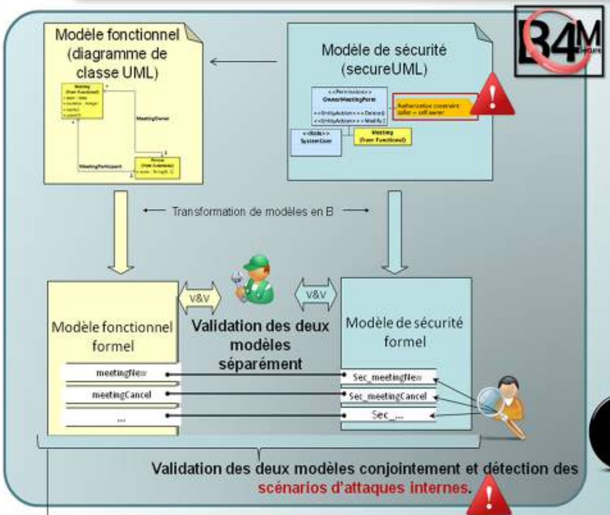
Sous la direction de : Yves LEDRU, Akram IDANI, Narjes BEN RAJEB

OBJECTIF

Validation formelle de politiques de contrôle d'accès sur deux plans:

- Vérification et génération automatique de cas d'usage normaux.
- Extraction des scénarios d'usage malicieux.

CONTEXTE



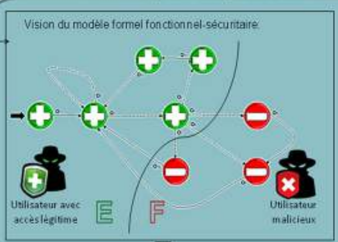
1) La validation de certaines propriétés de sécurité (comme les contraintes d'autorisation) dépend à la fois du modèle fonctionnel et du modèle de sécurité, d'où la nécessité de la prise en compte des deux modèles conjointement lors de la validation de la politique de sécurité.

2) L'extraction de scénarios d'attaques revient à trouver les chemins menant vers les états indésirables. Il s'agit, donc, d'un problème d'atteignabilité. Les techniques les plus connues pour résoudre ce genre de problème comme le model-checking souffrent du problème de l'explosion combinatoire de l'espace d'états quand le système est d'une grandeur importante.

Un **scénario d'attaque interne** est l'exécution d'une séquence d'opérations par un utilisateur malicieux **ayant le droit d'accès au système**, dans le but de s'octroyer de **nouveaux droits** et réaliser des opérations auxquelles il n'avait pas droit initialement.

SOLUTION

1. Recherche par la preuve dans un modèle symbolique



Prouver l'atteignabilité (ou non) de l'état indésirable (F) par certaines opérations (o_i) du système à partir de l'état (E).

$$\begin{cases} \exists x.E \wedge Pre(o_i) & \text{(a)} \\ \exists x.E \wedge Pre(o_i) \Rightarrow \neg[Action(o_i)] \neg F & \text{(b)} \end{cases}$$

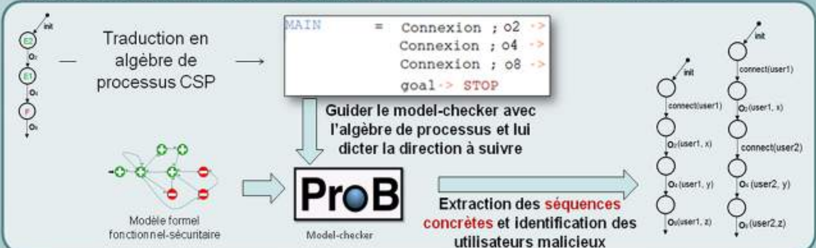
(a) Obligation de preuve vérifiant que l'opération (oi) est déclenchable à partir de l'état (E).

(b) Obligation de preuve vérifiant que l'opération (oi) permet d'atteindre l'état (F) à partir de l'état (E).

Réitérer le processus de preuve pour extraire tous les scénarios potentiels

séquences symboliques représentant des scénarios d'attaque potentiels extraites d'un modèle symbolique

2. Recherche par model-checking guidé dans le modèle concret



L'approche a permis l'extraction de scénarios d'attaque sur 3 cas d'étude. Les résultats sont très prometteurs pour l'expérimenter sur des cas réels.

PERSPECTIVES

- Application de l'approche de détection de scénarios d'attaque pour la génération et la validation de scénarios nominaux.
- Identification d'autres types d'attaques: élévation de privilège, déni de service, etc.
- Automatisation de l'approche et expérimentation en vraie grandeur.

REFERENCES

- Amira Radhouani, Akram Idani, Yves Ledru, Narjes Ben Rajeb. **Symbolic search of insider attack scenarios from a formal information system modeling.** LNCS Transactions on Petri Nets and Other Models of Concurrency (ToPNoC). Special issue of selected revised papers of FMS'2014 . (Accepté)
- Akram Idani, Yves Ledru, Amira Radhouani. **Modélisation graphique et validation formelle de politiques RBAC en Systèmes d'Information.** Ingénierie des Systèmes d'Information 19(6): 33-61 (2014). Numéro spécial « Sécurité des SI ».

VASCO

"Validation de Systèmes, Composants et Objets logiciels"

User-friendly digital 3D painting

Elisabeth Rousset, François Bérard, Mickaël Ortega

Ease 3D manipulation [1]

Problem

- How to map flat gestures to 3D manipulation ?
- Easy for translation, but state-of-the-art rotations are difficult

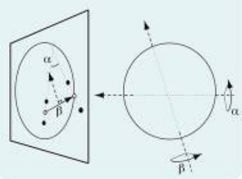
Physical world

Digital world

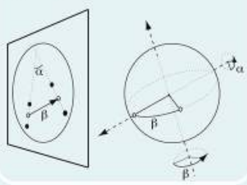
Contribution

- two new interaction techniques for 3D rotations
- techniques designed for multitouch screens
- application of the surjective concept to HCI

Two Axis Valuator +



Arcball +



Evaluation

Apparatus 16 participants performing each one 120 trials on a docking task, comparing our techniques against the state of the art

Results Our techniques are significantly faster and more precise than the state of the art, and are preferred by participants

Ease digital drawing

Problem

- Pen-enabled devices are used in 2D/3D computer graphics
- They are either graphics tablets or interactive pen displays
- Which type of device is the more user-friendly for a drawing task?

Graphics tablet

Interactive pen display

Hypothesis

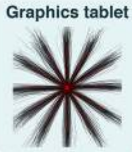
We know that graphics tablets introduce eye/hand dissociation and gain
We suppose that eye/hand dissociation and gain influence the orientation error and variability of strokes

Evaluation

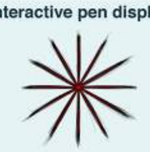
Apparatus 16 participants performing each one 900 trials on a drawing task, comparing graphics tablet versus interactive pen display

Results
- heterogenous gain increases strokes variability
- eye/hand dissociation increases strokes error, variability and completion time
- participants unanimously prefer interactive pen displays

Graphics tablet



Interactive pen display



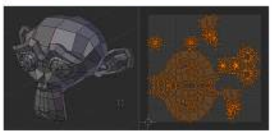
Conclusion Interactive pen displays are significantly easier to use for drawing tasks

Ease 3D painting

Problem

3D model texturing is performed either by unwrapping 3d meshes on flat surfaces for 2D painting or by direct 3D painting with projection algorithms

Unwrapp + 2D painting



High cognitive load to work out 2D projection in one's head

3D projection painting

Cognitive load for manipulation distracting from painting task

Contribution

A new set of interactions for easy 3d projection painting:
- designed for interactive pen displays
- manipulation is done by non-dominant hand via multitouch
- painting is done by dominant hand via stylus
- painting & manipulation may be simultaneous
- manipulation implies low cognitive load thus is not distracting

Evaluation

In progress...

IIHM

"Ingénierie de l'Interaction Homme-Machine"

Laboratoire d'Informatique de Grenoble

2015 LIG

[1] É. Rousset, F. Bérard, and M. Ortega, "Two-finger 3D rotations for novice users: surjective and integral interactions," Proc. Advanced Visual Interfaces, 2014.

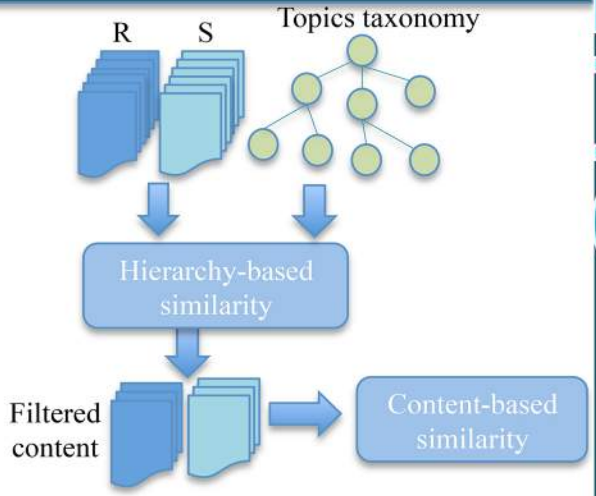
A Distributed Algorithm for Hierarchical Similarity Joins

Rafik Saad
LIG/SLIDE rafik.saad@imag.fr

Supervisors :
Sihem Amer-Yahia , Vincent Leroy

Motivation

- Many datasets are organized in hierarchical structures (taxonomies, ontologies).
- Joins are essential and expensive operations in data management.
- Objective : Process hierarchy-based similarity joins (given similarity thresholds)
- Applications :
  - Finding articles with similar topics
  - Grouping Yago entities.



Challenges and Approach

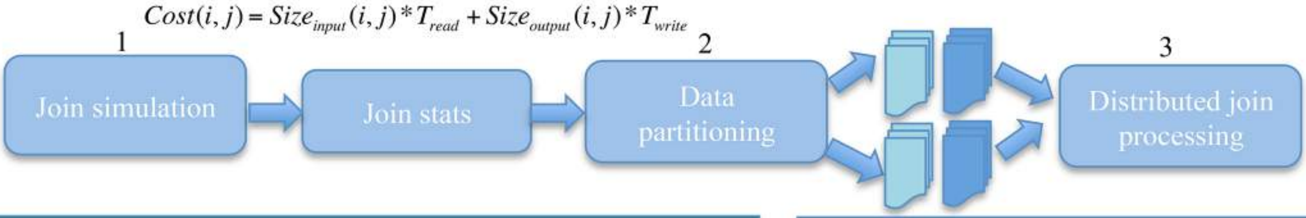
- Challenge : high I/O cost.
- Partitioning algorithm :
  - Use statistics from a simulation phase
  - Estimate partitions costs, using hardware benchmarks
  - Greedily minimize partitions unbalance

$$Stats(i) = \{N^{buffer}, Size_{input}, Size_{output}\}$$

$$Cost(i,j) = Size_{input}(i,j) * T_{read} + Size_{output}(i,j) * T_{write}$$

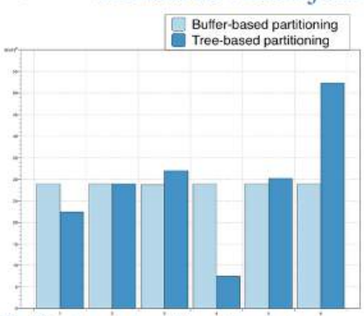
- Approach : partitioned join processing, buffer-based data partitioning scheme.
- Buffer-management strategy : next-usage based, size-aware

$$w(a) = size(a) * \Delta t$$

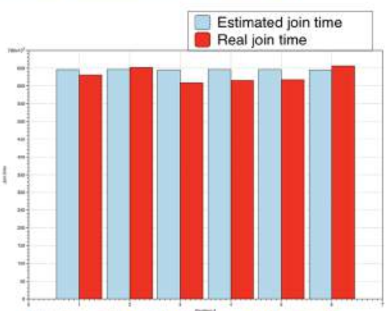


Experiments

- Datasets : BioASQ, Dbpedia, Yago
- Measures : total join time, partitions balance



Buffer-based vs Tree-based partitioning



Costs estimation vs real costs

Conclusion

- **Contribution** : a scalable distributed algorithm for hierarchical similarity joins processing
- **Future work** : adapt the algorithm to streaming data, n-ary joins

SLIDE

"ScaLable Information Discovery and Exploitation"

Modèles de construction de géovisualisations pour la représentation des dynamiques des processus spatiaux : Application au risque d'inondation impactant le système ferroviaire

Contexte et Problématique

Cette thèse se situe dans le domaine de la géomatique et s'intéresse au processus de construction cartographique et à l'utilisation de **données spatio-temporelles, hétérogènes**, **peu structurées** et **imparfaites**, comme le sont les données historiques. Elle se déroule dans le cadre d'une collaboration SNCF Infra / LIG Steamer qui vise à mieux comprendre la dynamique des catastrophes naturelles, notamment les inondations, ayant impactées le système ferroviaire (infrastructure et circulations des trains).

Objectifs :

1) Définir des **modes de visualisation spatio-temporelle**, adaptés aux spécificités des données historiques et aux utilisateurs, facilitant l'extraction de connaissances sur les dynamiques liant phénomènes spatiaux et système ferroviaire.
2) Automatiser le processus de **construction de géovisualisation** en prenant en compte les caractéristiques des données et les besoins des utilisateurs.

Démarche méthodologique

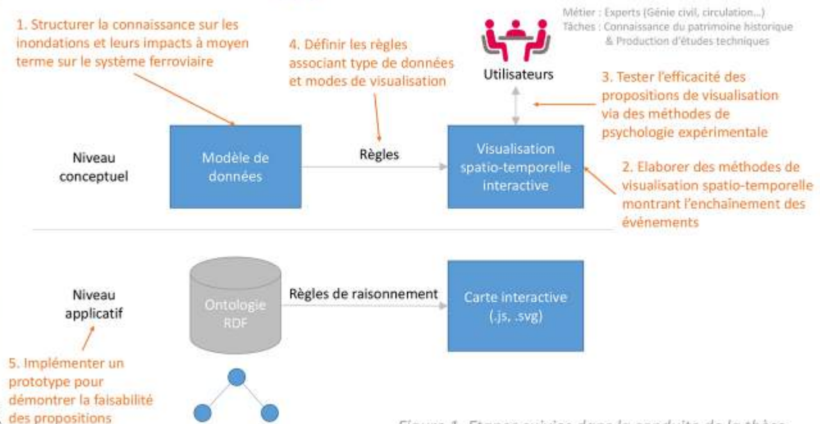


Figure 1. Etapes suivies dans la conduite de la thèse

Structuration de la démarche

Basée sur une méthode d'ingénierie logicielle classique, notre démarche présente deux originalités :

- L'**évaluation cognitive des visualisations**: le retour des utilisateurs sur leurs préférences d'interface ne suffit pas, car la visualisation préférée n'est pas toujours celle qui donne les meilleurs résultats d'analyse [voir Kinkeldey et al. (2014) How to Assess Visual Communication of Uncertainty? *The Cartographic Journal*] ;
- La proposition de **règles** permettant de définir automatiquement le **mode de visualisation spatio-temporelle le plus adapté**, en fonction des **caractéristiques spécifiques du jeu de données** à visualiser (ex: selon le type de temporalités, la densité spatiale des évènements...) et du **profil métier de l'utilisateur** (accent mis sur les objets d'intérêt du système ferroviaire – cf. Ontologie IDISFER)

Contributions à mi-parcours

Contribution 1: L'ontologie IDISFER

Ontologie IDISFER : Impact Des Inondations sur le Système FERroviaire

Verrous scientifiques :

- Articulation entre réseau de transport et son exploitation (système complexe),
- Descriptionsdes impacts sur les objets de ce système,
- Relations méréologiques et de causalité entre événements.

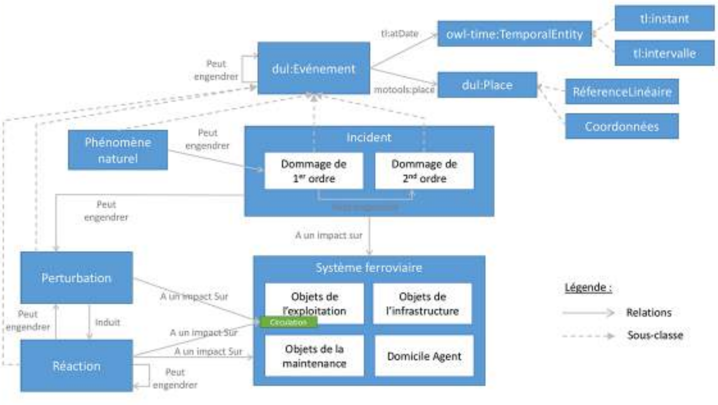


Figure 2. Modèle simplifié de l'ontologie IDISFER

Contribution 2: Des formes de visualisation temporelle

Verrous scientifiques :

- Comment **visualiser les relations de causalité** entre événements (effet domino), d'une manière qui facilite l'analyse,
- Quelles **méthodes de visualisation du temps** dans les cartes sont efficaces pour réaliser des tâches d'**analyse** complexes,
- Comment **restituer l'imperfection des données** historiques pour qu'elle ne soit plus un obstacle à leur utilisation.

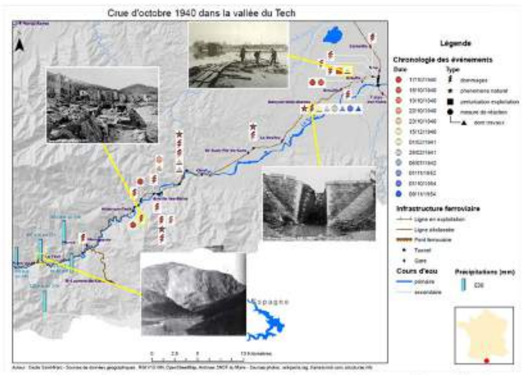


Figure 3. Proposition de carte chronologique multimédia des événements lors de la crue de 1940

Conclusions et perspectives

Bilans des travaux :

- La mise en place d'une **méthode** de collecte et de **capitalisation des données d'archives** ;
- La construction du **modèle de données IDISFER** (graph RDF) ;
- L'identification des **besoins des utilisateurs**, les **spécifications** d'un outil de visualisation des inondations passées impactant le système ferroviaire ;
- La proposition de premières **géovisualisations d'événements**.

Perspectives :

- Produire des **modes de visualisations** adaptés aux **temporalités** et aux **imprécisions** des données historiques ;
- Conduire les **expérimentations** pour tester l'adaptation des modes de visualisation aux tâches des utilisateurs, interpréter les résultats ;
- Mettre en place les **règles de choix** du mode de géovisualisation le plus adapté au contexte.

A terme, la thèse devra fournir des approches cartographiques adaptées aux données, efficaces et utilisables dans un contexte opérationnel pour résoudre des tâches complexes. Les résultats seront généralisés à d'autres types de risques et d'autres réseaux.

STEAMER "Spatio-temporal information, adaptability, multimedia and knowledge representation"

Laboratoire d'Informatique de Grenoble 2015 LIG

Cécile Saint-Marc – Thèse CIFRE SNCF – LIG/Steamer
Janvier 2014 – Décembre 2016
cecile.saint-marc@imag.fr – http://membres-lig.imag.fr/saintmar/

Directrice de thèse : Paule-Annick Davoine
Co-directrice : Marlène Villanova-Oliver

Ontology alignment and data interlinking evolution on the web of data

Adam Sanchez-Ayte
(adam.sanchez-ayte@inria.fr)
Supervisors: Jérôme Euzenat, Jérôme David

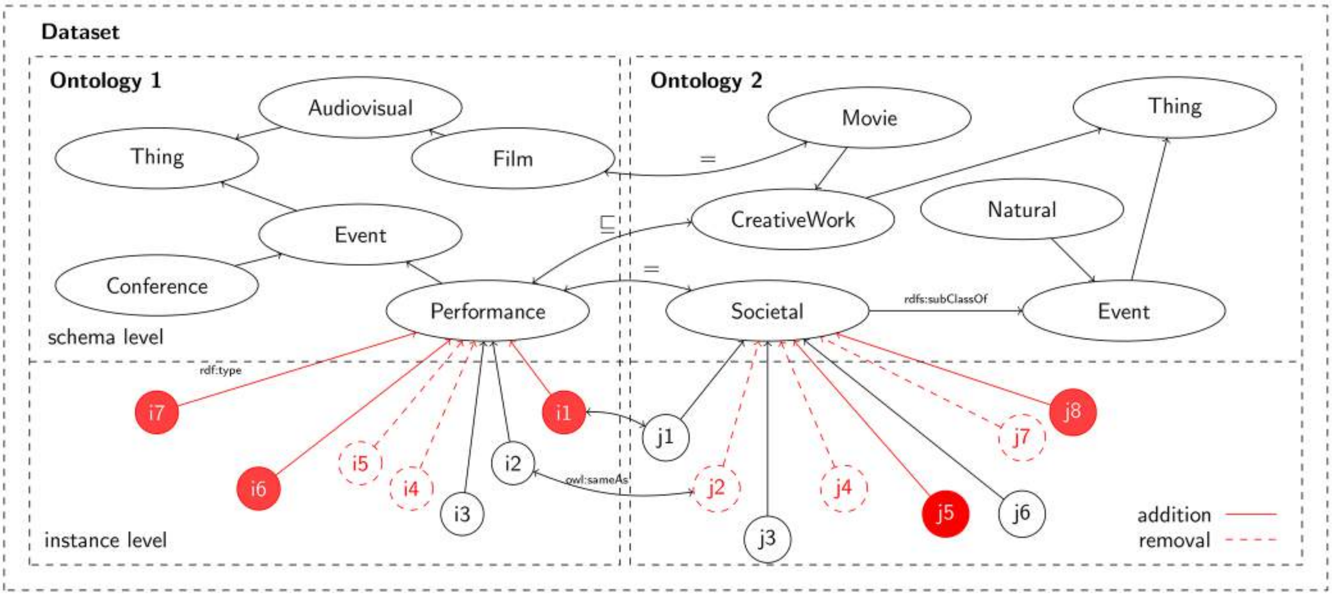
Goal: Provide automatic and reliable method to maintain alignments and links when ontologies and data change.

Context: The semantic web has to evolve.

If ontologies and data evolve, then ontology alignments and links between data have to evolve too.

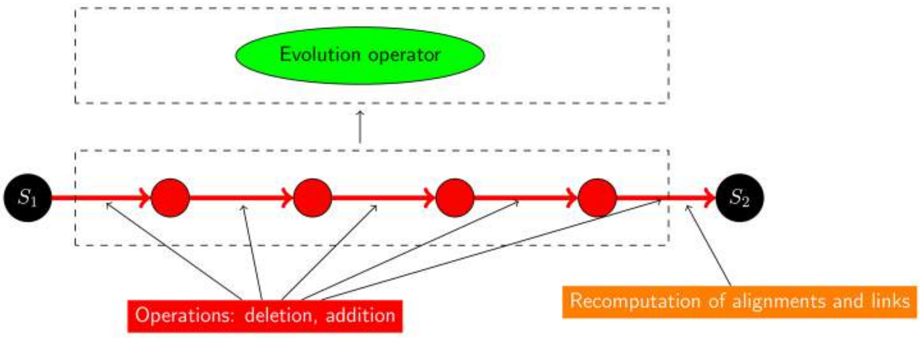
Problem: How to seamlessly maintain its coherence?

Given a set of changes to be performed in a dataset, develop a method to maintain alignments and links.



Proposed approach

- Capturing dependencies between inferred data.
- Designing change operations to revise them when data evolves.
- Characterizing such operations with rational postulates. (as in Belief Revision [1])



First experiments with instance-based matching and linkkey extraction.

References

[1] Giorgos Flouris. On belief change in ontology evolution. *AI Commun.*, 19(4):395–397, 2006.

EXMO

"Computer mediated exchange of structured knowledge"

Socio-Affective Intelligence for Robot

Yuko Sasa

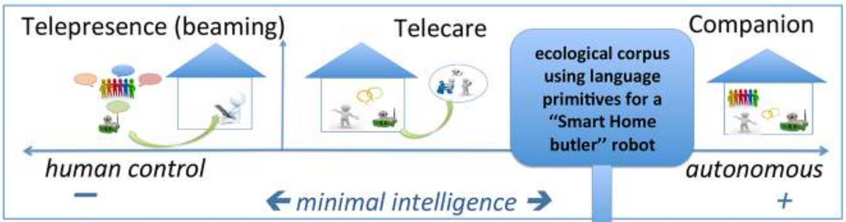
Véronique Aubergé (LIG-CNRS) – Feng Gang (Gipsa-lab) – Yoshinori Sagisaka (Waseda University)

Context and Hypothesis

"Socio-affective glue " as key of dynamic dialog
no glue without social role – no social role without glue

Similar communication functions carried by language and more primitive speech cues as pure socio-affective prosody events

Understanding interaction process through a damaged communication situation of socio-affective relation (e.g typically elderly)



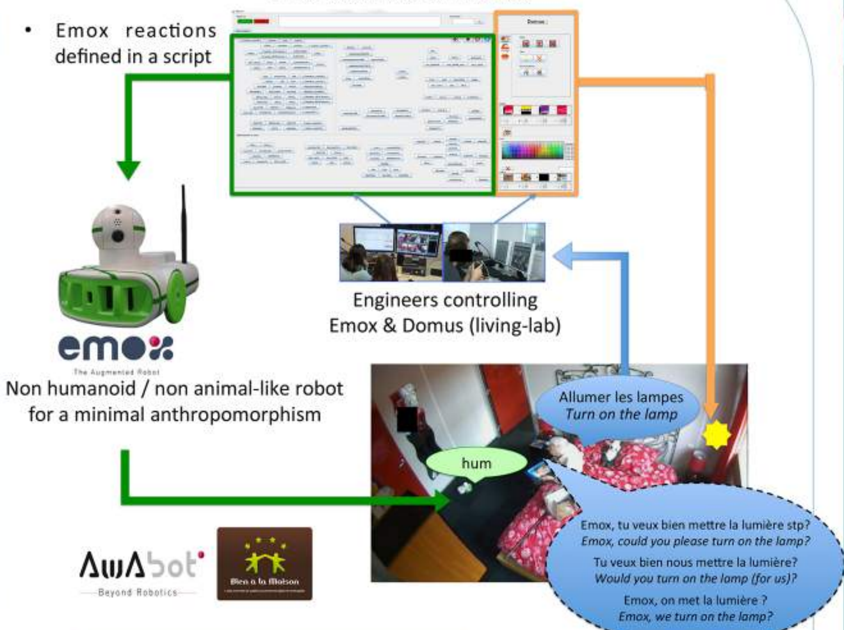
Graduated Prosody Levels

1) No speech → 2) Pure prosodic mouth noises supposed to be the "glue's" tools → 3) Lexicons with supposed "glue" prosody → 4) Commands imitations with supposed "glue" prosody

Experiments

Scenario

1) **Visit of the Smart Home** : an engineer welcomes an elderly, his caregiver and a student who recruited the subjects for a gerontechnology study dealing with the effect of personal objects on people habituation of technologized place.
2) **Fake emergency call from the control room:** the caregiver and the student have to leave the Smart Home without having the time to see the robot.
3) **Introducing the butler Emox :** the engineer introduces the robot a list of 30 commands that can be adressed to robot but which has to be all tested once.
4) **Emox-Elderly Interactions:** The elderly test all the commands and can accomodate the Smart Home with his objects while waiting
5) **Return of the caregiver then the student** : each one asking how is working the Smart Home
6) **Debriefing** : Asking what elderly thought about the Smart Home (and not the robot) – Real aim of this study and questions about the robot.



Elderly Emox Expressions Corpus

- Spontaneous but controlled natural interactions
- 30 voice commands list given to the subjects
- 10 frail and socio-isolated elderly over 75 years old recorded (30-50 subjects expected)
- Muldimodal speech data

Labelling and auto-annotation

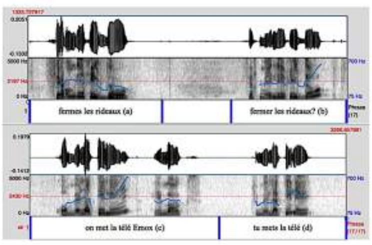
Elderly review and annotate themself their vocal and gestural behaviors in guided by the expeimenters. Labels will be perceively validated

Domus action, visual and vocal cues of the robot and the subjects are marked (dialog rythm and interpersonal synchrony - [Chetouani, 2012])

Results & Perspectives

Elderly commands changes along interactions with Emox

a) Imposed syntax before glue processing – reading like attitude
b) Imposed syntax with increasing positive glue prosody (systematic F0 arise at the end of the sentences + breathy voice).
c) Various paraphrasing (used in synergy with "we") = globally high F0 and a great arise at the end of the sentences.
d) Increasing use of prosodic focalizations
e) Non commands productions to guide or to recover error without entering in a dialogue process
f) Comments without answer expectation, politness, envionment description, cheering, caring



Evaluate''socio-affective glue''degree key points through a global visualization of micro-dialog events

Dialogue SASI
Socio-Affective Speech Interaction

Communication dynamic changes analysis to elaborate a socio-affective dialog bootstrap to train for interaction, including :
Expressive speech synthesis
Ecological speech recognition
=> A long-life incremental machine learning system to structure interaction for a robot in a specific social role

GETALP
"Groupe d'Étude en Traduction Automatique/Traitement Automatisé des Langues et de la Parole"

Laboratoire d'Informatique de Grenoble

2015 LIG

Acknowledgments :
Nicolas Bonnefond, Brigitte Meillon, Florian Nebout, Sandrine Tie Bi, Tim Robert, Jonathan Rey-Gorrez, Leandra Antunes, Adrien Schwartz, Romain Magnani, Frédéric Aman, Leslie Guillaume, Maxence Girard-Rivier, Liliya Tsvetanova, Cécile Cottier, Noémie Lagier, Amaury Nègre, Sybille Caffiau

ANR 11-LABX-0025-01

Time Series Centroid Estimation under Weighted and Kernel Dynamic Time Warping

Saeid SOHEILY KHAH   Superviseurs: Ahlame DOUZAL, Eric GAUSSIER

Université Joseph Fourier, LIG-AMA, Saeid.Soheily@Imag.Fr

Definition

Time series is a kind of **sequence** data

- an ordered sort of elements
- order criteria : **time**

Time Series analysis are involved in: **Signal processing** and **Pattern recognition**, **Load curve prediction**, **User behavior analysis** and etc.

Motivation

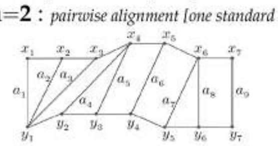
Estimating a centroid of a set of time series is involved in many data mining and machine learning processes as:

- **Summarizing** a set of time series
- **Extracting** temporal prototype
- **Clustering** time series

Problem Statement

Challenging Question : centering more than two times series under temporal warping

- **n=2** : *pairwise alignment [one standard way]*



	1	2	...	9
x	x_1	x_2	...	x_7
y	y_1	y_1	...	y_7
c	avg(x_1, y_1)	avg(x_2, y_1)	...	avg(x_7, y_7)

- **n>2** : *need to handle the problem of multiple temporal alignment*

Alternative Approaches :

- Dynamic Programming
  - *Search by* **dynamic programming** *for optimal path* $[O(T^N)]$
- Progressive Approach
  - **Combining pairwise** *alignment through different strategies*
- Iterative Approach
  - *Progressive approach +* **repeatedly** *realign all the time series*

Drawbacks :

- **Heuristics** approaches
- High time and space **complexity**
- Don't guarantee to reach **global optimal solution**
- **Fail** for time series that share only local characteristics

Objectives

Formalize the multiple time series averaging problem as an optimization problem

- Propose an **optimal solution** for barycenter estimation
- Estimate a **weight vector**

Solution

Propose a new averaging process under:

- **Weighted DTW** metric
- **Kernel DTW** metric

Formalization Problem

Determine centroid **c** and its weight vector **w** that optimize the inertia criteria :

- $\arg\min_{\mathbf{c},\mathbf{w}} \sum_{i=1}^{N} \text{WDTW}(\mathbf{x}_i, (\mathbf{c}, \mathbf{w}))$
- $\arg\max_{\mathbf{c},\mathbf{w}} \sum_{i=1}^{N} \text{KDTW}(\mathbf{x}_i, (\mathbf{c}, \mathbf{w}))$
  - $X = \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ *: set of time series*
  - $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})$ *: a time series of length T*
  - $\mathbf{c} = (c_1, c_2, ..., c_T)$ *: centroid of* X
  - $\mathbf{w} = (w_1, w_2, ..., w_T)$ *: weighting vector* $[\sum_{t=1}^{T} w_t = 1, w_t \geq 0, \forall t]$

Centroid Estimation

Def. 1) $\text{WDTW}(\mathbf{x}_i, (\mathbf{c}, \mathbf{w})) = \min_{\pi_i \in \mathbb{A}_i} \underbrace{\frac{1}{|\pi_i|} \sum_{(t',t) \in \pi_i} f(w_t)\, \varphi(x_{it'}, c_t)}_{C(\pi_i)}$

- **Aim** : To **minimize** the cost of warping path(π_i)
  - ϕ: an Euclidean norm
  - $f(w)$: warping constraint function (behaves opposite to w)

Def. 2) $\text{KDTW}(\mathbf{x}_i, (\mathbf{c}, \mathbf{w})) = \max_{\pi_i \in \mathbb{A}_i} \underbrace{\frac{1}{|\pi_i|} \sum_{(t',t) \in \pi_i} f(w_t)\, \kappa(x_{it'}, c_t)}_{C_\kappa(\pi_i)}$

- **Aim** : To **maximize** the cost of warping path(π_i)
  - k *: a kernel similarity between aligned observation*
  - $f(w)$ *: an increasing function (behaves similar to* w *)*

Solution I - WDTW

- Centroid Estimation(**c**)

$$c_t = \frac{1}{|\mathbf{N}_t|} \sum_{i=1}^{N} \sum_{(t,t') \in \pi_i^*} x_{it'}$$

- π_i^* *: optimal alignment between* $\mathbf{x}_i$ *and* $\mathbf{c}$
- $\mathbf{N}_t = \{t' | (t,t') \pi_i^*\}, \forall i \in \{1,...,N\}$ *: the neighborhoods of t*

- Weight Estimation(**w**) $[f(x) = x^{-\alpha}]$

$$w_t = \frac{A_t^{\frac{1}{1+\alpha}}}{\sum_{t=1}^{T} A_t^{\frac{1}{1+\alpha}}} \qquad A_t = \sum_{i=1}^{N} \frac{1}{|\pi_i^*|} \sum_{(t,t') \in \pi_i^*} \varphi(x_{it'}, c_t)$$

- A_t *: average divergence on the neighborhood* $\mathbf{N}_t$ *of time stamp t*

Solution II - KDTW

- Centroid Estimation(**c**) : For K taken as the **Gaussian** kernel,we adopt a gradient ascent approach:

$$c_t^{(p+1)} = c_t^{(p)} + \eta^{(p)} \frac{\partial L}{\partial c_t^{(p)}} \qquad [\,\eta^0 = 1\,,\ \eta^{P+1} = \eta^p / p\,]$$

$$\frac{\partial L}{\partial c_t} = \sum_{i=1}^{N} \frac{1}{|\pi_i^*|} \sum_{(t,t') \in \pi_i^*} f(w_t)\, \frac{2}{\sigma^2} (x_{it'} - c_t)\, e^{(-\frac{\|x_{it'} - c_t\|^2}{\sigma^2})}$$

- Weight Estimation(**w**) $[f(x) = x^{\alpha}]$

$$w_t = \frac{A_t^{\frac{1}{1-\alpha}}}{\sum_{t=1}^{T} A_t^{\frac{1}{1-\alpha}}} \qquad A_t = \sum_{i=1}^{N} \frac{1}{|\pi_i^*|} \sum_{(t,t') \in \pi_i^*} \kappa(x_{it'}, c_t)$$

- A_t *: average similarity on the neighborhood* $\mathbf{N}_t$ *of time stamp t*

Results

DATASET	ALTERNATIVE APPROACHES				PROPOSED APPROACHES	
	RC-DTW	HC-DTW	AM-DTW	IAM-DTW	TSC-WDTW	TSC-KDTW
CBF	38.3%	42.3%	-61.3%	32.1%	**81.7%**	54.3%
CC	39.8%	28.6%	6.8%	13.6%	70.5%	**72.3%**
DIGITS	26.1%	79.5%	77.6%	82.2%	**87.6%**	82.9%
TRAJ	67.1%	87.7%	85.2%	90.6%	**93.1%**	90.6%
BME	34.9%	43.1%	-11.8%	59.4%	**91.2%**	83.8%
UMD	-6.5%	51.1%	-56.2%	48.8%	**89.7%**	73.3%
SPIRAL	59.8%	64.4%	64.2%	65.8%	**72.8%**	63.6%
SPIRAL2	61.4%	66.3%	9.3%	9.8%	**98.9%**	83.7%
CONSSEASON	84.1%	70.5%	4.6%	21.4%	**95.5%**	92.7%

Table 1: Comparison of Inertia Reduction Rate

DATASET	ALTERNATIVE APPROACHES						PROPOSED APPROACHES		
	RC-DTW		HC-DTW		IAM-DTW		TSC-WDTW		TSC-KDTW
	length	time	length	time	length	time	time	length	time
CBF	8283	392.32	35042	9999.89	128	42.91	**61.27**	128	147.83
CC	992	4.15	1677	12.73	60	6.46	**3.23**	60	15.63
DIGITS	313	0.52	530	1.09	85	0.52	**0.47**	85	1.99
TRAJ	33	0.06	29	0.06	20	0.03	**0.02**	20	0.10
BME	2027	5.46	2781	11.92	95	3.93	**3.48**	95	7.02
UMD	2729	10.32	4280	28.87	121	4.75	**4.15**	121	12.31
SPIRAL	660	1.62	1122	3.33	95	1.19	**0.83**	95	3.80
SPIRAL2	1699	16.13	9030	269.93	300	34.84	**15.74**	300	153.22
CONSSEASON	5741	77.10	32706	3680.81	144	29.79	**21.85**	144	66.98

Table 2: Comparison of Time/Space complexity

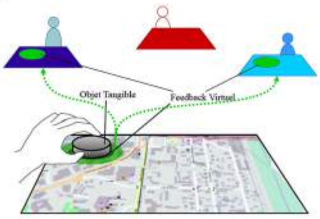
Forthcoming Research

Estimate the barycenter in input space by optimizing criteria in feature space in Kernels

"Analyse de données, Modélisation et Apprentissage automatique" AMA

Lauren Thévin
lauren.thevin@imag.fr
LIG CNRS, Université de Grenoble - UMR 5217 - Laboratoire LIG - Maison Jean Kuntzmann
110 av. de la Chimie - Domaine Universitaire de Saint-Martin-d'Hères
BP 53 - 38041 Grenoble cedex 9 – France
En collaboration avec Institut des Risques Majeurs

Décisions collectives en environnements interactifs et collaboratifs complexes : application à la gestion de crise



Nous présentons une approche pour la conception d'un système multi-agent de préparation des acteurs communaux à la gestion de crises, dans le contexte d'une interaction via des tables tangibles distribuées. Notre point focal est celui des normes qui régissent ces interventions. Le rôle du système est d'analyser la validité des interventions et de gérer l'interaction grâce à l'interprétation et à l'inférence via des normes.

POINTS CLES

Hypothèse : conscience organisationnelle partagée nécessaire à la collaboration

Contexte applicatif : exercice d'entraînement à la gestion de crise supporté par des interactions tangibles.

Solution : SMA (Système Multi-Agent) normatif (1) pour la régulation de l'activité humaine et (2) pour la mise en œuvre des interactions distantes et fournir des retours informés aux acteurs selon les 3 facettes Production, Communication, Coordination

PROBLEMATIQUE

Contexte :

- Multiplicité et hétérogénéité des acteurs et des organisations
- Coordination des interventions des acteurs et de leurs réactions aux aléas : répartition et non duplication des efforts, gestion des ressources
- Distribution géographique des acteurs

Verrous :

- Plusieurs points de vue organisationnels
- Normes situées dans un contexte (environnement et organisation)
- Coordination à distance

Exemple de solutions déjà existantes :

Médiation mise en place à travers l'interaction Homme-Machine dans un contexte de collecticiel :

- modèles fonctionnels (proches de la réalisation) : ARCH, Cameleon
- modèles conceptuels (proches de la conception) : PAC, AMF-C
- modèles explicitant les deux aspects (de manière néanmoins hétérogènes) PAC-Amodeus, PAC*, AMF-hybride

TRAVAUX ENGAGES/RESULTATS

Approche :

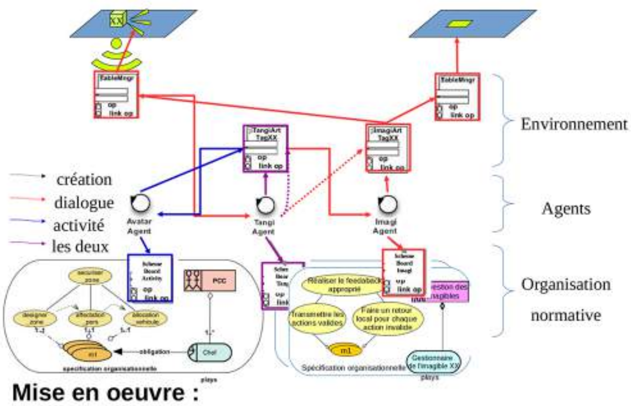
SMA normatif selon le framework JaCaMo pour répondre à des situations d'interaction complexes (suivi de l'action humaine et de l'interaction H-M-H) :

- Interaction mixte riche
- Système hybride humain/artificiel.

Principes :

- Interaction mixte : tangibles, imagibles, digibles
- SMA normatif hybride : JaCaMo S = <A,E,I,O>

Architecture :



Mise en oeuvre :

- Implémentation
- Définition d'un scénario écologique (inondation)
- Modélisation/formalisation de normes

Publications :

Publications acceptées :

L. Thévin, F. Badeig, J. Dugdale, O. Boissier, C. Garbay. Rendre tangible des normes de collaboration par un système multi-agent normatif. Congrès RJCIA, juin 2014

L. Thévin, F. Badeig, J. Dugdale, O. Boissier, C. Garbay. Un système multi-agent normatif pour la collaboration et l'interaction mixte. Congrès JFSMA, octobre 2014

Maiquel De Brito, Lauren Thevin, Catherine Garbay, Olivier Boissier, and Jomi F. Hübner. Situated Artificial Institution to support advanced regulation in the field of Crisis Management, PAAMS 2015

Maiquel De Brito, Lauren Thevin, Catherine Garbay, Olivier Boissier, and Jomi F. Hübner. Situated Regulation on a Crisis Management Collaboration Platform, Demo at PAAMS 2015

Publications sous presse (révisions mineures) :

L. Thévin, F. Badeig, J. Dugdale, O. Boissier, C. Garbay. Un Système Multi-Agent normatif hybride pour l'interaction mixte : application à la gestion de crises. Revue RIA, 2015

Publication soumise :

Maiquel De Brito, Lauren Thevin, Catherine Garbay, Olivier Boissier, et Jomi F. Hübner. Institution artificielle située pour porter la régulation dans le domaine de la gestion de crise, JFSMA 15).

MAGMA AMA
"Modélisation d'agents autonomes en univers multi-agents"
"Analyse de données, Modélisation et Apprentissage automatique"

Tablet-Based Orchestration of Collaborative Activities in Classrooms

Patrick Wang, Pierre Tchounikine and Matthieu Quignard
patrick.wang@imag.fr

Introduction

- Technologies only recently entered classrooms.
- Computer-Supported Collaborative Learning (CSCL) is about designing software applications to support and scaffold collaborative activities among learners.
- It is difficult for teachers to prepare and conduct such activities.
- They need information, in real-time, regarding the development of the activity and tools to manage and/or modify it.

Objectives:

- Identify the difficulties teachers might encounter during collaborative activities in classrooms.
- Model, design and propose a software architecture dedicated to the orchestration of collaborative activities:
    - Answer teachers' needs.
    - Easy to use.
    - Robust and flexible to adapt to classrooms.

Current State of Work

"Orchestration refers to how a teacher manages in real-time multi-layered activities in a multi-constraints context." (Dillenbourg, 2013)

Figure 1. Teacher monitoring the students' progressions.

Figure 2. Teacher helps a group of students during the collective phase.

Interests of tablets over computers (Alvarez, Brown & Nussbaum, 2011):

- Provide mobility.
- Encourage face-to-face interactions.

Figure 1 and 2 illustrates how tablets can allow teachers to monitor students' and groups' progressions and help them.

In this case (Fig. 3), the teacher needs:

- information about the students' and groups' tasks;
- tools to prepare the individual and collective phase;
- tools to edit the collective phase.

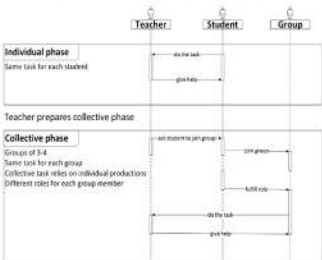


Figure 3. Model of a studied collaborative activity.

- To support monitoring for orchestration, we propose, for each phase of the activity, quantitative and qualitative information regarding the students' and groups' advancements.
- This allows for coarse-grained and finer-grained perception of how the activity unfolds.
- This allows teachers to identify and help students that are stuck or manage heterogeneity in students' speeds.

Results and Discussion

Research questions:

- Are the monitoring tools easy to use?
- Does the information provided allow teachers to satisfactorily monitor their students, groups and classes?

Experimentations and results:

- Four experimentations conducted in four different primary school classes (two 4th grade classes and two 5th grade classes, 4 teachers, 100 students).
- A paper-based individual activity has been transposed into a tablet-based CSCL activity (Wang, Tchounikine & Quignard, to be published).
- In each class, the experiment was conducted twice as the first half worked on the collaborative activity while the other half was given another task, and vice-versa.
- Students did not encounter difficulties while doing the collaborative activity.
- Each teacher acknowledged that they were comfortable with using the application after the first experiment, implying that the tools are easy to use and that tablets are a relevant support.
- Despite using different strategies, each teacher affirmed that they could satisfactorily monitor each student and group, and intervene when needed.

Discussion:

- Although teachers regarded the tools as useful, the relatively few amount of students working on the collaborative activity might have an impact on the teachers' perceptions.
- The simple aspect of the tools let teachers spontaneously adopt different monitoring strategies. User-friendliness is a primordial feature when studying the teachers' acceptance of such technologies (Roschelle, Dimitriadis & Hoppe, 2013).

Future Work

- Design and implement tools that allow teachers to edit the parameters of the collaborative activity in the light of the monitoring information.
- Design and implement tools that facilitates the preparation of the collective phase
- Upgrade the software architecture to permit flexibility in the management of the activity.
- Explore the genericity of the application with different collaborative activities.

References

Alvarez, C., Brown, C., & Nussbaum, M. (2011). Comparative study of netbooks and tablet PCs for fostering face-to-face collaborative learning. Computers in Human Behavior, 27(2), 834-844.

Dillenbourg, P. (2013). Design for classroom orchestration. Computers & Education, 69, 485-492.

Roschelle, J., Dimitriadis, Y., & Hoppe, U. (2013). Classroom orchestration: Synthesis. Computers & Education, 69, 523-526.

Wang, P., Tchounikine, P., & Quignard, M. (to be published). Orchestration Challenges Raised by Transposing a Paper-Based Individual Activity into a Tablet-Based CSCL Activity: An Example. 11th International Conference on Computer-Supported Collaborative Learning

METAH
"Modèles et Technologies pour l'Apprentissage Humain"

Software Transactional Memory with Autonomic Management Techniques

Naweiluo Zhou, Gwenael Delaval, Eric Rutten and Jean-Francois Mehaut

Grenoble University, France; French Institute for Research in Computer Science and Automation (INRIA), France
Naweiluo.Zhou@inria.fr; gwenael.delaval@inria.fr
eric.rutten@inria.fr; Jean-Francois Mehaut:jean-francois.mehaut@imag.fr

"Compiler Optimization and Run-time SystEms"

CORSE

1. Introduction

Concurrent programs need to manage time trade-off between synchronization and computing. A high concurrency level may decline computing time but increase synchronization cost among threads. More threads may be more conflict! The traditional way to handle synchronization problems is through implementing locks. However locks suffer from:

- likelihood of deadlocks.
- vulnerability to failures, faults...

2. Platform

- Hardware: SMP machine (24 Intel core, 64G memory) & SMP machine (192 Intel cores, 800G memory)
- Software:
  - STM/control system: tinySTM, Heptagon
  - programming languages: C/C++, Heptagon
  - benchmarks: Eigenbench, STAMP

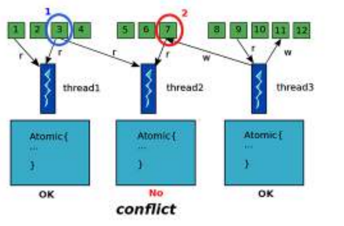
3. State of the Art*

- Bottleneck of resource management for HPC systems
- Computing systems require to adapt its behaviour dynamically
- Still a new idea involving control theory in TM systems

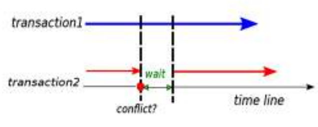
4. Transactional Memory

Addressing synchronization issues through transactions:

- An alternative data structure support to lock-based memory structure.
- A new way for synchronization on multicore processors, accesses to data are wrapped in transactions.
- Transactions may be committed or aborted.



(a)



(b)

Figure 1: An example of transactions access to memory. In case1 thread1 and thread2 both read object3; In case2 thread2 reads from object7 and thread3 writes to object7. The conflict in case2 can lead to wrong results

Figure 2: Backoff contention manager policy. When a conflict happens, a transaction chooses to wait and then continues

Three main contention manager policies

- Aggressive: when a transaction detects a conflict, it kills other transactions
- Suicide: when a transaction detects a conflict, it kills itself
- Backoff: when a conflict happens, a transaction waits and then continues

5. Autonomic Management

- Control techniques for supervising computing systems
- Design feedback control loops to manage computing systems

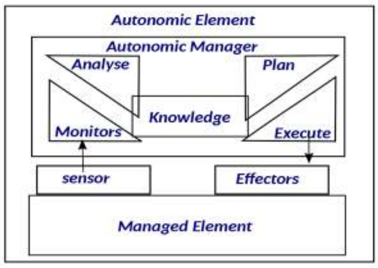
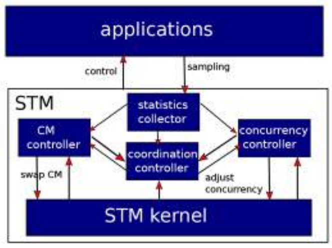


Figure 3: A feedback control loop. It incorporates 4 elements: managed element, autonomic manager (includes monitor, analyse, plan, knowledge and execute), sensor (collect information) and effector (carry out changes to managed element

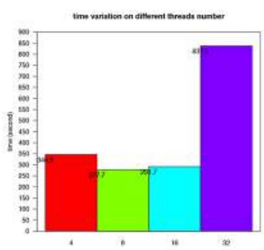
6. Approaches

- Introduce autonomic control techniques into STM system
- Design feedback control loops to automatic the STM control
- Coordinate control loops to optimize performance

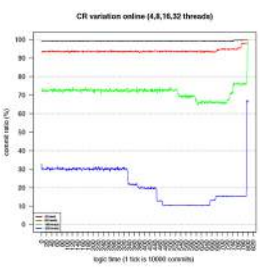
Figure 4: How feedback control loops can be implemented in TM systems



7. Results



(a)



(b)

Figure 5: Time variation on different concurrency level

Figure 6: Commit ratio variation on different concurrency level at runtime

***Reference:**

– 1. An, Xin. "High Level Design and Control of Adaptive Multiprocessor Systems-on-Chip." PhD diss., Grenoble University, 2013.
2. Castro, Marcio Bastos. "Improving the Performance of Transactional Memory Applications on Multicores: A Machine Learning-Based Approach. " PhD diss., Universit de Grenoble, 2012.
3. D. Rughetti, P. Di Sanzo, B. Ciciani, and F. Quaglia. Machine learning-based self-adjusting concurrency in software transactional memory systems. In Modeling, Analysis Simulation of Computer and Telecommunication Systems (MASCOTS), 2012 IEEE 20th International Symposium on, pages 278–285, Aug 2012.
4. M. Ansari, C. Kotselidis, K. Jarvis, M. Luján, C. Kirkham, and I. Watson. Advanced concurrency control for transactional memory using transaction commit rate. In Proceedings of the 14th International Euro-Par Conference on Parallel Processing, Euro-Par '08, pages 719–728, Berlin, Heidelberg, 2008. Springer-Verlag.

Inria

Laboratoire d'Informatique de Grenoble

2015 LIG

Acknowledgments:
This thesis is funded by the French National Research Agency (ANR) and PERSYVAL Lab.