Control Theory for Computing Systems

Application to big-data cloud services & location privacy protection

Sophie Cerf

Supervised by Nicolas Marchand and Bogdan Robu

16.05.2019 - PhD Defense

gipsa-lab

Communauté UNIVERSITÉ Grenoble Alpes

Computing Systems Context

- Growing in number and complexity
- Used by all actors of the society
- Service vs. Platform

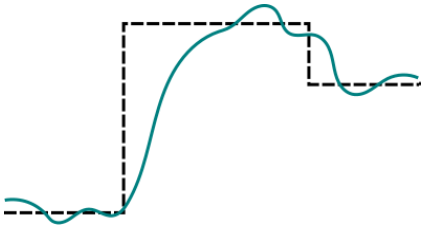
- Objectives: Automated software adaptation to guarantee performance, availability, security, etc.

control
reliability
dependability
robustness
performance
security
guarantee
privacy
adaptation
availability
costs

Control Theory Approach

Automated tuning of a system to guarantee objectives

- Properties
  - Evolution through time
  - Observable and tunable system
- Objectives
  - Outputs performance
  - Stability
- Modeling
- Control
- Evaluation

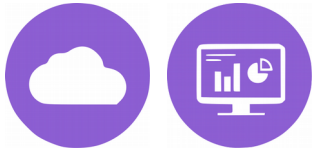


Application Use-Cases

- Location Privacy

  - Automated database protection [SRDS'17], [IEEE TDSC'19]
  - Dynamic formulation, modeling and control [DAIS'18] [CCTA'18]

- BigData Cloud Services



  - Adaptive Robust Control [IFAC WC'17]
  - Optimal Cost-aware Control [CDC'16]

- Learning Algorithms

  - Robust Learning on Unreliable Data [NIPS'18] [DSN'19]
    - 5-months internship @IBM Research, Zurich
  - Feedback-based Training of Neural Networks [CCTA'19]

Outlines

I. Background
- Computing System Context
- Control Theory Approach
- Application use-cases

II. Location Privacy

III. BigData Cloud Services

IV. Conclusions

Outlines

I. Background

II. Location Privacy

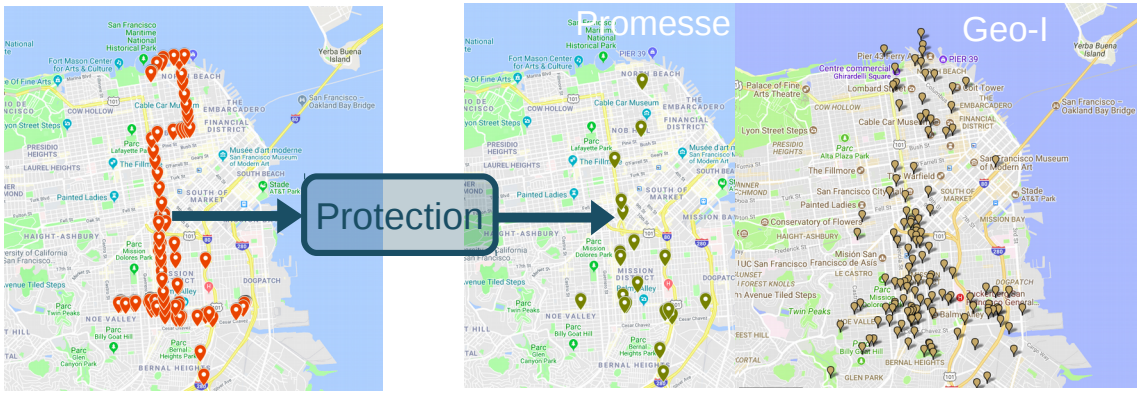
- Introduction, Related Works and Objectives
- Control Problem Formulation
- Modeling: developpement and validation
- Control: formulation and evaluation

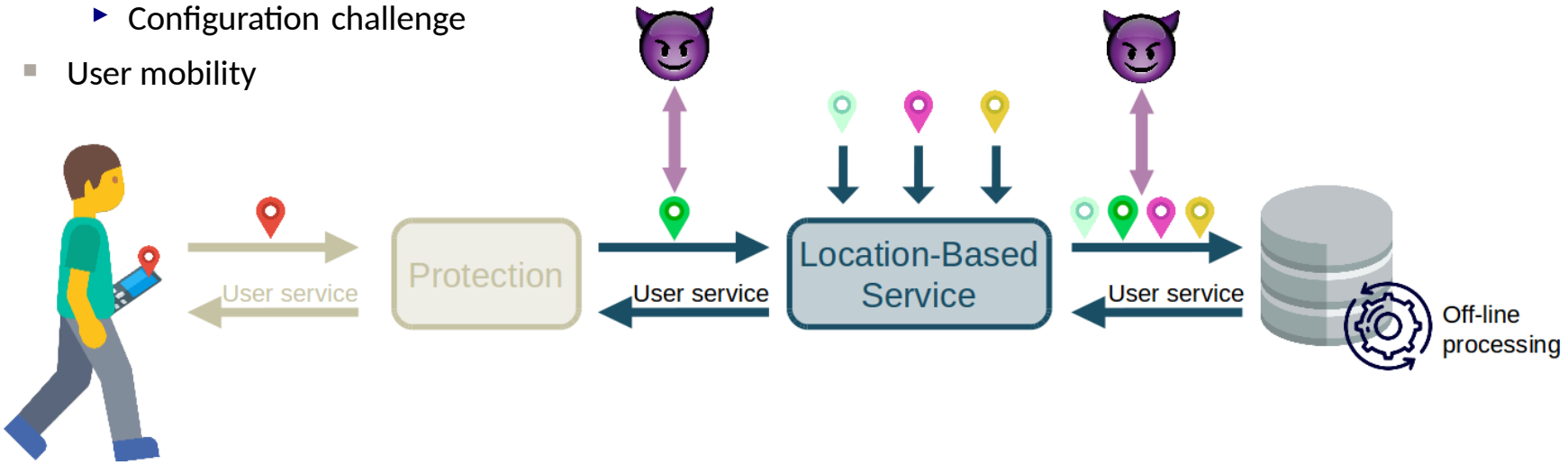
III. BigData Cloud Services

IV. Conclusions

Introduction on Location Privacy

- Scenario
  - User perpective : dynamic location broadcating
  - Location-Based Service perspective : mobility databases
  - Privacy
    - Motivation : leaks, regulations
  - Location Privacy Protection Mechanism (LPPM)
    - Configuration challenge
  - User mobility





Related Work

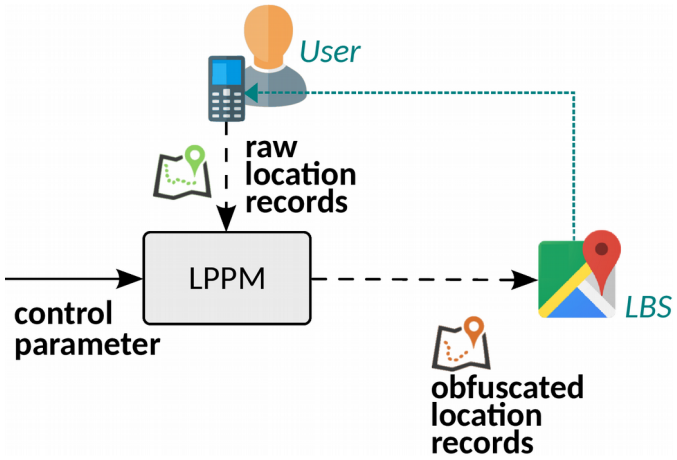
- State of the Art
  - (Agir, 2014) iteratively modify the configuration to meet the privacy objective
    - Limits: Computing intensive, no utility
  - (Chatzikokolakis, 2015) adapts Geo-I's parameter to the density of the area
    - Limits: not objective driven, no utility
  - (Primault, 2016) iteratively evaluates the privacy and utility for refining configuration parameters
    - Limits: Computing intensive
- Open Challenges
  - No performance guarantees
  - Privacy is not dynamic, i.e. no history is taken into account
  - No robustness regarding user's mobility

Objectives & Main Results

- Objectives
  - Automatic guarantees of privacy and utility for users and services through LPPM configuration, with robustness regarding user's movements.
- Main Results
  - Automatic objective-based LPPM choice and configuration for mobility datasets
    - [SRDS'17] [IEEE TDSC]
    - *PULP* framework (Matlab)
  - Online utility-aware modeling and control of location privacy
    - [DAIS'18] [CCTA'18]
    - *dynULP* framework (Matlab)

Control Problem Formulation

- System
  - Mobile device user broadcasting her location dynamically to receive a service
  - Use of an online LPPM
    - Geo-I (differential privacy)
- Inputs
  - Amount of spatial noise added to the location
  - User mobility
- Outputs
  - Privacy
    - Diameter of the largest Point of Interest from the last time window
  - Utility
    - Spatial distortion

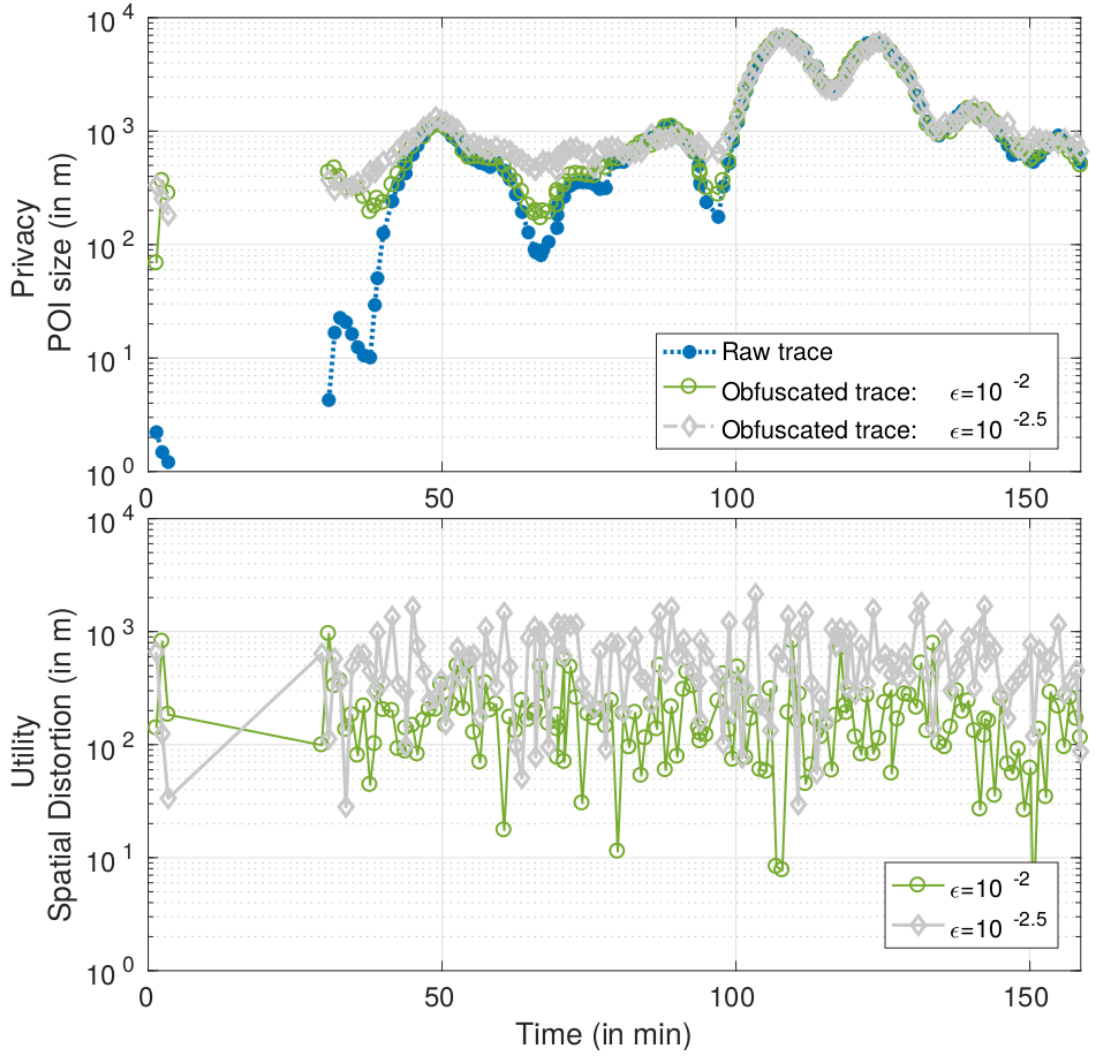


$$\rightarrow priv(k) = 2 \times median\left(dist[l(t), l_c(k)]\right)$$

$$l_c(k) = \frac{1}{T} \sum_{t=k-T}^{k} l(t)$$

$$\rightarrow util(k) = dist[l(t), l'(k)]$$

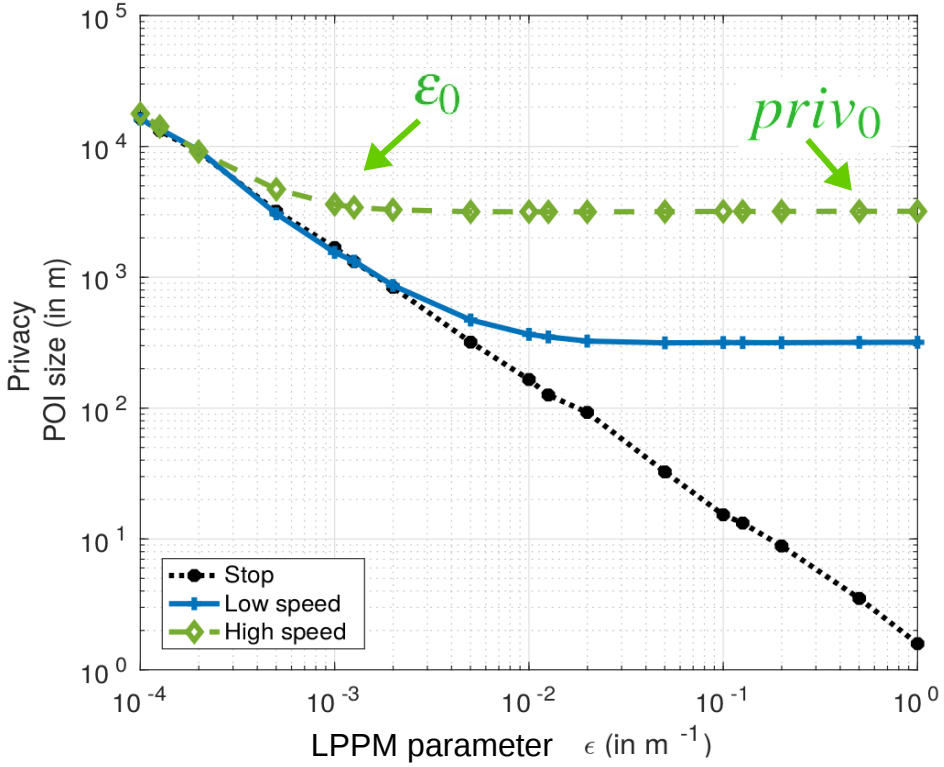
Motivation

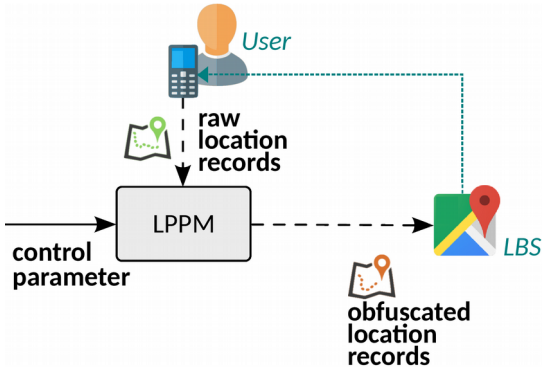

Privacy
POI size (in m)
Raw trace
Obfuscated trace: ε=10 -2
Obfuscated trace: ε=10 -2.5
Utility
Spatial Distortion (in m)
ε=10 -2
ε=10 -2.5
Time (in min)

Modeling

Static – Dynamic – Validation

- Methodology:
  - Set the input parameter and measure the output privacy, for different user movements





Linear part:

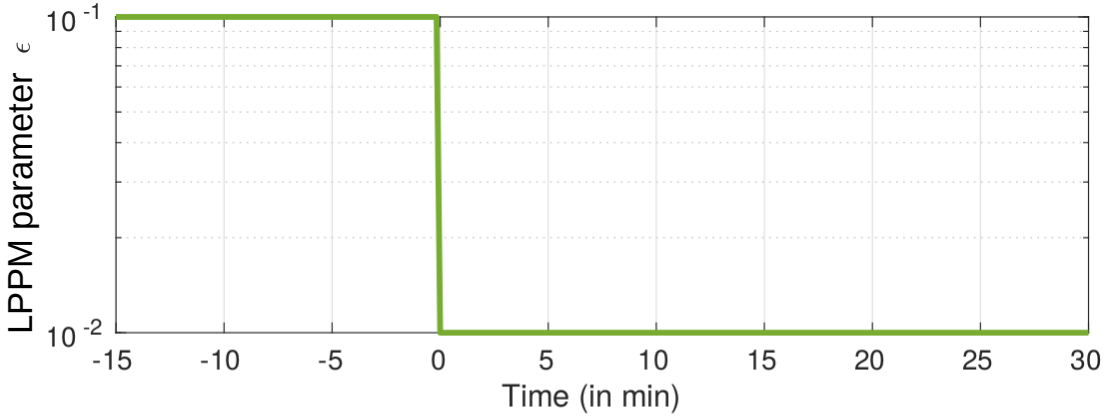
$$log(priv) = a\, log(\epsilon) + b.$$

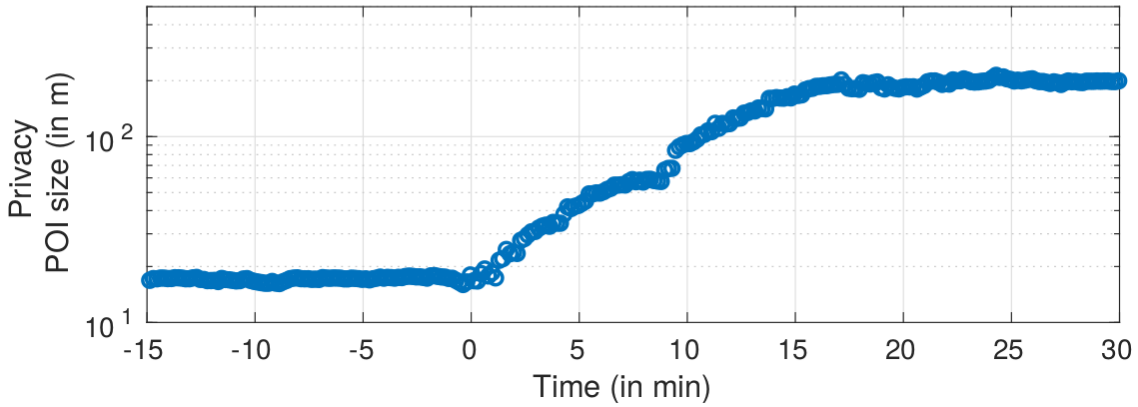
Linearisation:

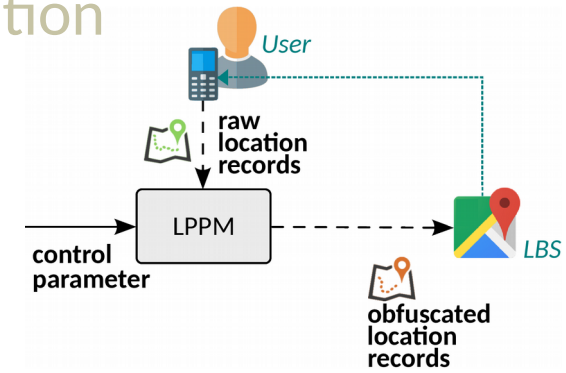
$$\Delta\varepsilon = log(\varepsilon) - log(\varepsilon_0),$$
$$\Delta Priv = log(priv) - log(priv_0).$$

Modeling

Static – Dynamic – Validation





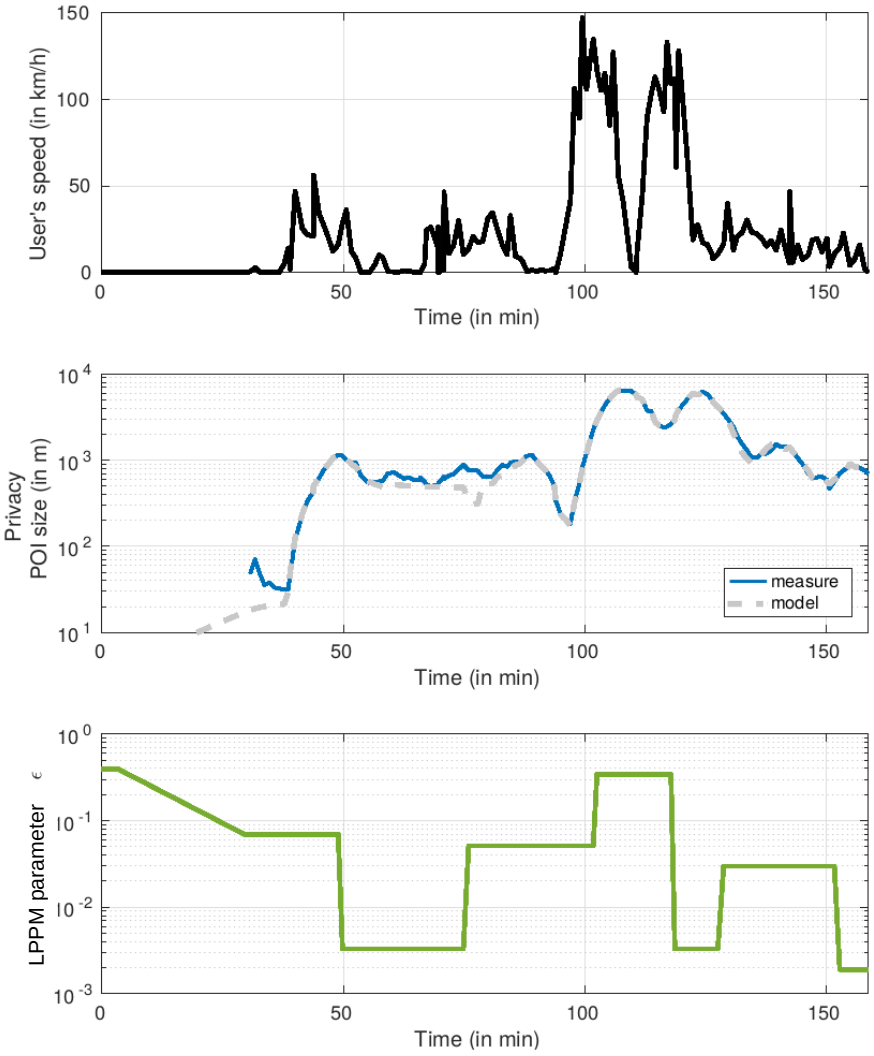


In the linear part $(\epsilon < \epsilon_0)$, without perturbation:

$$H(s) = \frac{\Delta Priv(s)}{\Delta \epsilon(s)} = \frac{K}{1 + \tau s}$$

Modeling

Static – Dynamic – Validation



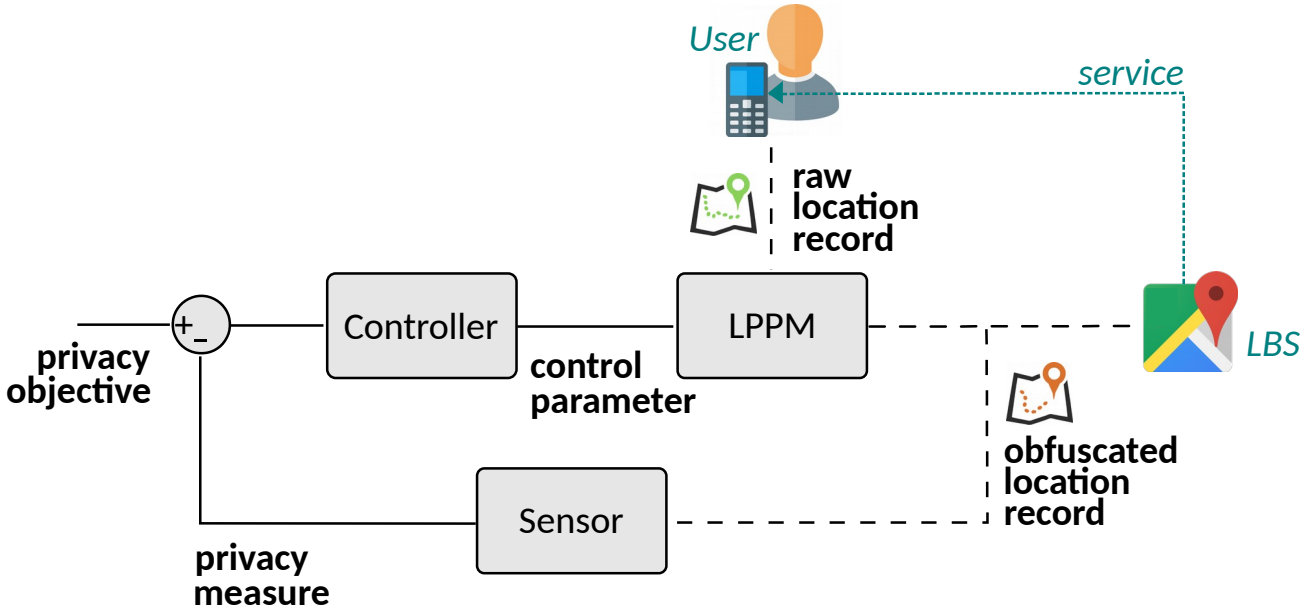
- Input scenarios
  - Mobility trace
    - comprehensive synthetic trace
    - real-life record
  - LPPM parametrization
- Modeling performance indicators
  - Normalized log squared

$$\frac{1}{|K|}\sum_{k\in K}\frac{(log(priv_{model}(k)) - log(priv_{measure}(k)))^2}{log(priv_{measure}(k))}$$

Control

Formulation – Evaluation

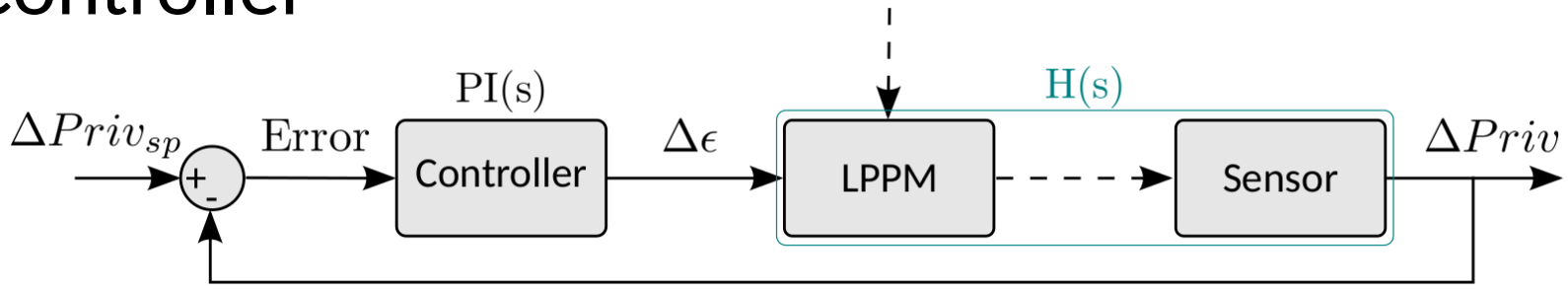
- Objectives
  - Robust privacy reference tracking
  - Increase utility when possible



Control

Formulation – Evaluation

- Controller



  - Linear formulation

$$\Delta Priv_{sp} = log(priv_{sp}) - log(priv_0).$$

  - Anti Windup

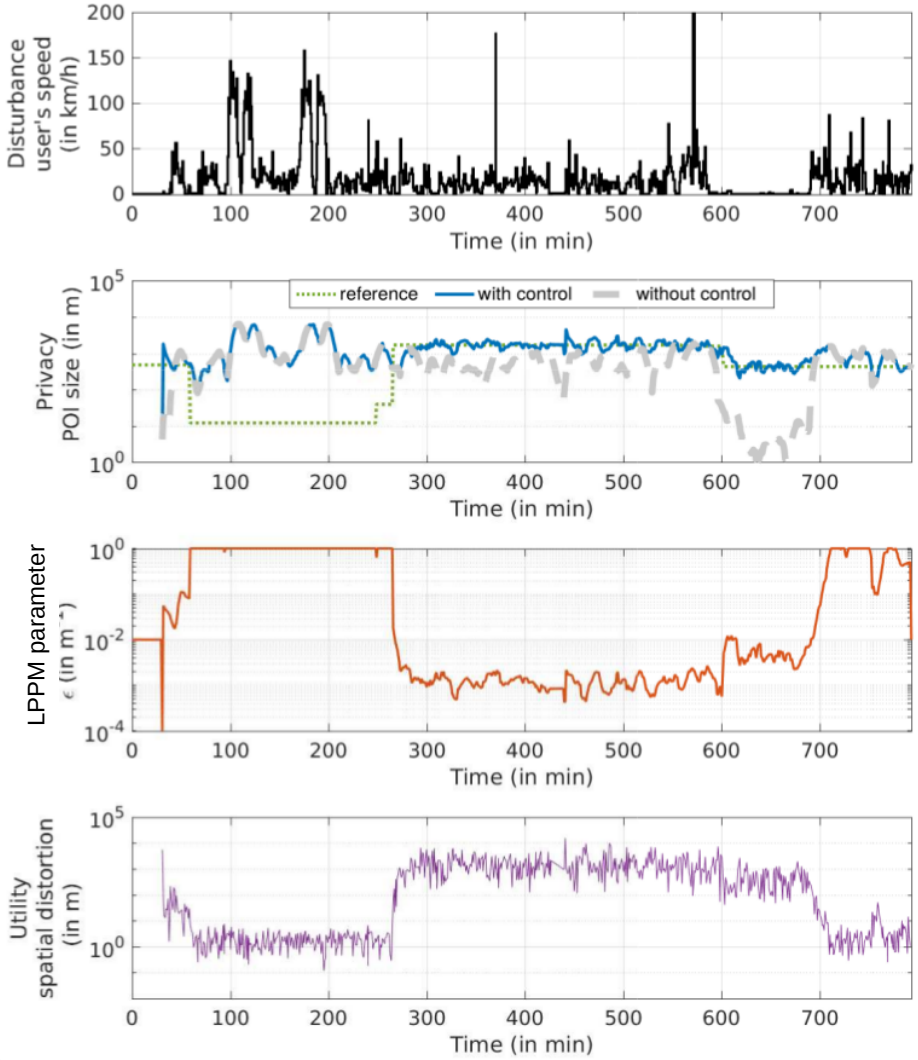
$$\varepsilon(t_i) = min(max(\varepsilon_{PI}(t_i), \varepsilon_{min}), \varepsilon_{max}).$$

  - Proportionnel Integral

$$PI(s) = \frac{\Delta\varepsilon(s)}{\Delta Priv_{sp}(s) - \Delta Priv(s)} = \frac{K_I}{s} + K_P.$$

Control

Formulation – Evaluation



- Input scenarios
  - Real mobility trace
  - Privacy specification
- Performance indicators
  - Privacy regulation
    - normalized log square overshoot

$$\frac{1}{|K|}\sum_{k\in K}\frac{(max[log(priv_{sp}(k))-log(priv_{measure}(k)),0])^2}{log(priv_{sp}(k))}$$

  - Utility preservation
    - median, 99th percentile
  - Execution time

Outlines

I. Background

II. Location Privacy

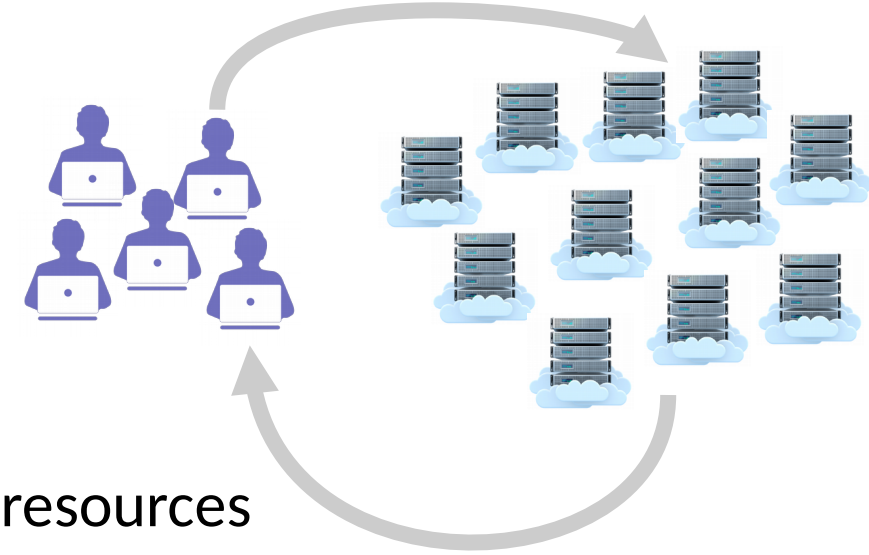
III. BigData Cloud Services

- Introduction and Objectives
- Related Work: Problem Formulation & Modeling
- Robust Control
- Cost-aware Control

IV. Conclusions

Introduction on Cloud Services

- Scenario
  - Bigdata processing framework Hadoop/MapReduce
    - Can realize many tasks
    - Widely used in the industry
    - Master-slave architechture
  - Cloud based : shared, scalable resources

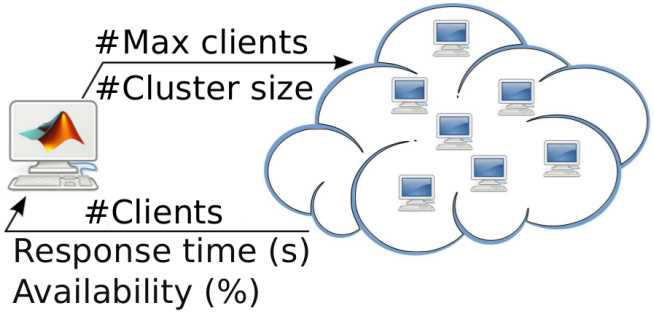


- Challenges
  - Overcome the dynamic workload
  - Deal with the cloud variability
  - Not intrusive solution

Related Works

- State of the Art
  - Static, Reactive, Predictive techniques
    - (Ali-Eldin, 2012) (Nguyen, 2013)
  - Control theory : Problem formulation, modeling, controllers, experimental setup
    - (Berekmeri, 2016) (Bekcheva, 2018)
- Open Challenges
  - Robustness to environment and system's changes
  - Include cost considerations

amazon web services™ | EC2

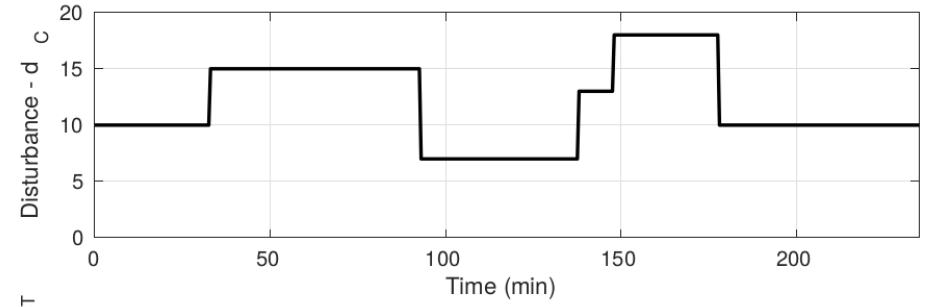


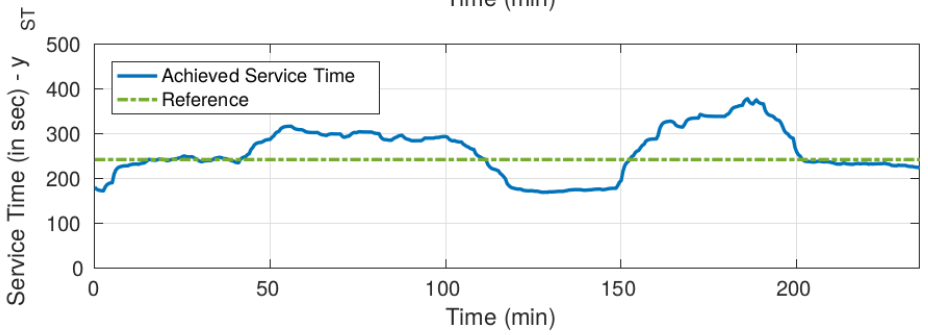
Objectives and Main Results

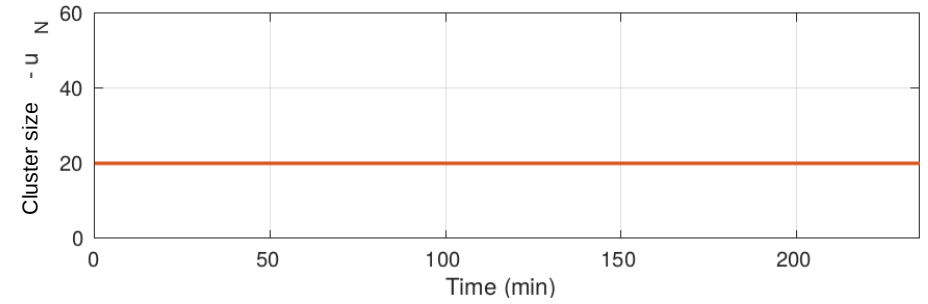
- Objectives
  - Ensure desired **performance** and **availability** of MapReduce while being **robust** to the changes in the application and its environment; in a **cost-aware** way
- Main Results
  - Adaptive controller robust to environment and system's changes
    - [IFAC'17]
    - Experimental setup (Linux bash, Matlab)
  - Multi-objectives cost-aware optimal controller
    - [CDC'16]
    - Simulation setup (Matlab Simulink)

Cloud Control Formulation

- System
  - Hadoop/MapReduce
  - Grid5000 cloud
- Outputs
  - Service time
    - Averaged execution time of jobs over the last 15 min
  - Availability
- Inputs
  - Number of nodes in the cluster
  - Maximum number of admitted clients
  - Clients workload
    - uncontrollable
- Open loop behavior

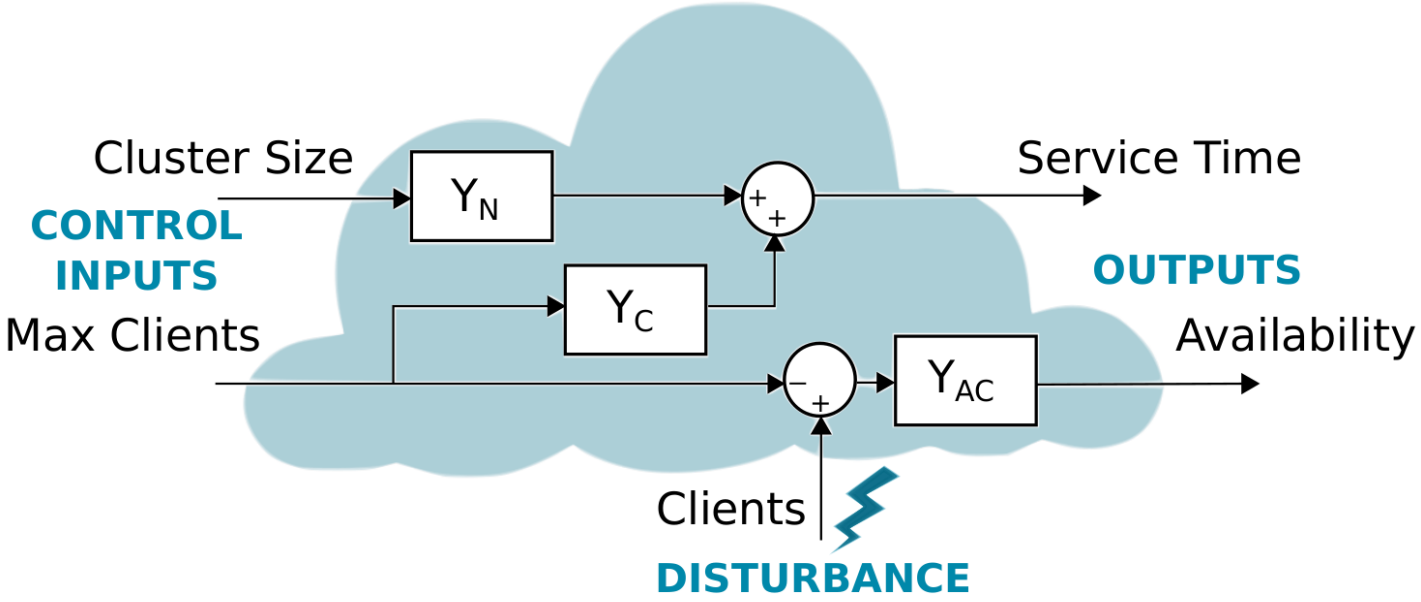






Modeling

State of the Art



- First order with delay, discrete time models

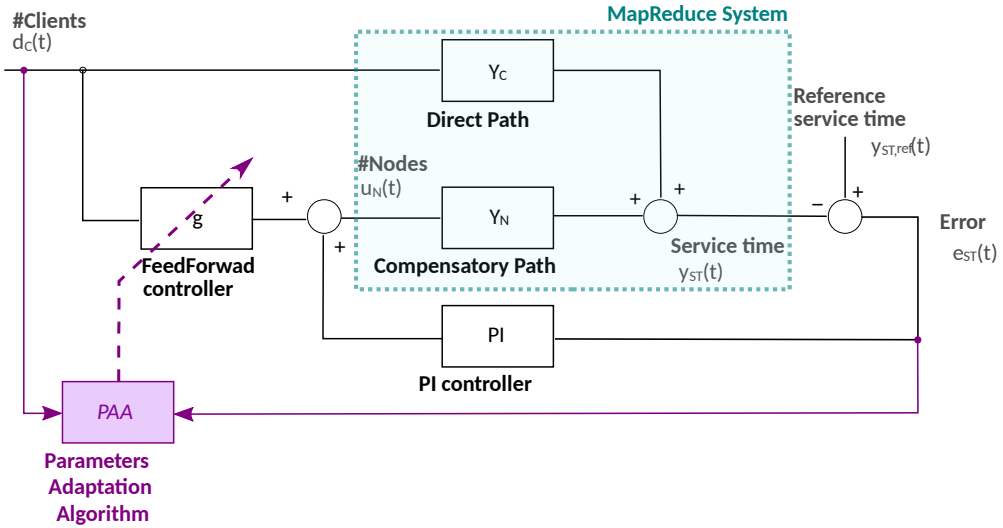
$$Y_C(q^{-1}) = \frac{b_C q^{-1}}{1 + a_C q^{-1}} q^{-r_C}, \quad Y_N(q^{-1}) = \frac{b_N q^{-1}}{1 + a_N q^{-1}} q^{-r_N}, \quad Y_{AC}(q^{-1}) = \frac{b_{AC} q^{-1}}{1 + a_{AC} q^{-1}}$$

- Parameters found by identification

Robust Cloud Control

Formulation – Stability – Performance

- Objectives
  - Stable service time
  - Robustness to workload changes and to system variability
- Formulation
  - PI feedback
    - steady state convergence
  - Static Feedforward
    - disturbance rejection
  - Feedforward parameter adaptation
    - robustness
- Preliminary hypothesis
  - the effect of the PI controller can be neglected
  - The two transfer functions have the same dynamics
  - and their delays are known and equal



Robust Cloud Control

Formulation – Stability – Performance

Adaptative Feedfoward gain:

$$\hat{g}(t+1) = \hat{g}(t) + \alpha x(t+1)e(t+1)$$

where

$$x(t) = \frac{sgn(b_N).q^{-(r_N+1)}}{1+\hat{a}_N q^{-1}} d(t)$$

THEOREM

$$lim_{t\to\infty} e(t+1)\,0$$

provided that $H'(z^{-1}) = \frac{1+\hat{a}_N z^{-1}}{1+a_N z^{-1}}$

is a strictly positive real transfer function

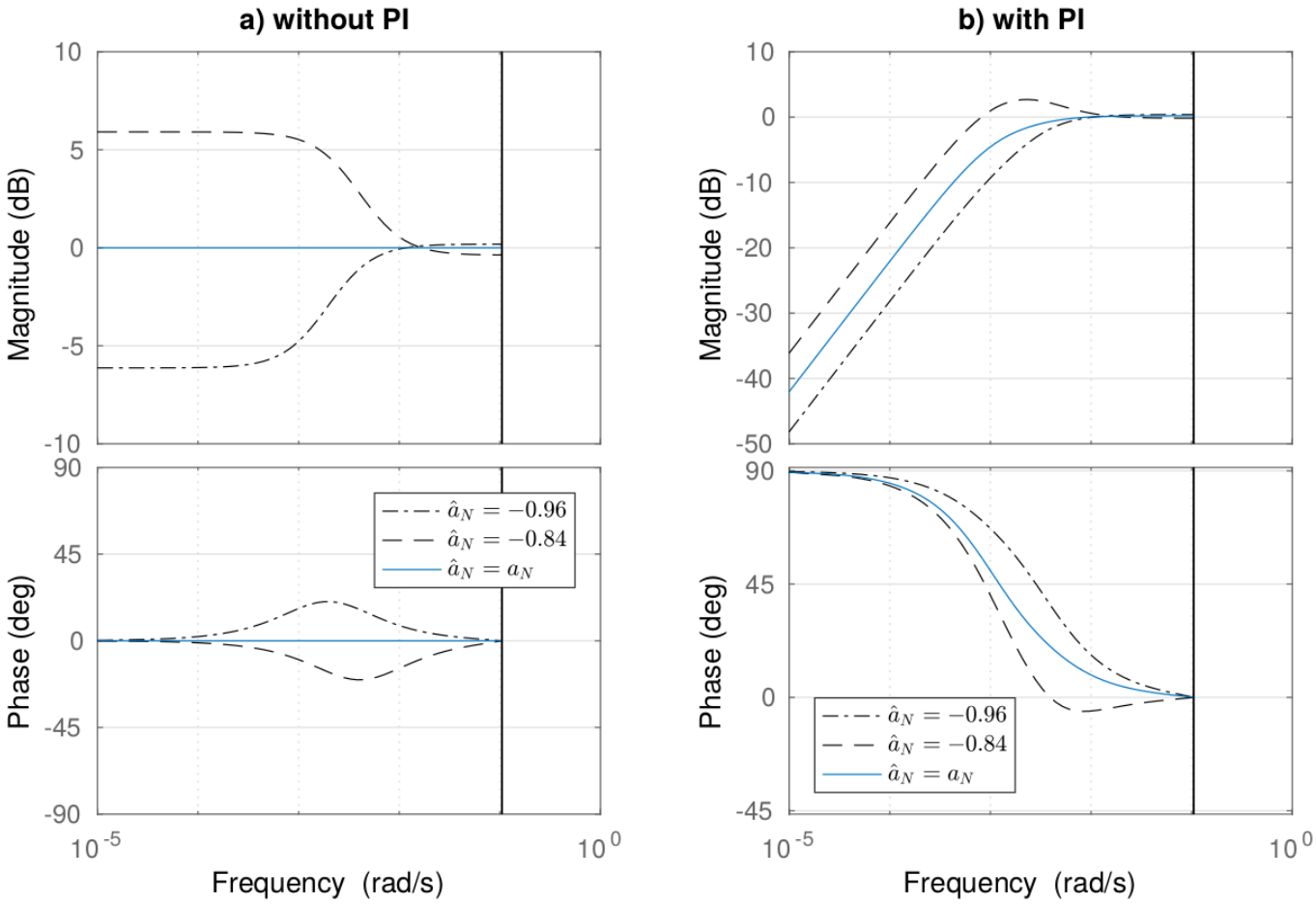
HINT OF PROOF

$$e(t+1) = \frac{|b_N|(1+\hat{a}_N q^{-1})}{1+a_N q^{-1}}(g^* - \hat{g}(t+1))x(t)$$

Optimal value of the feedforward controller

Robust Cloud Control

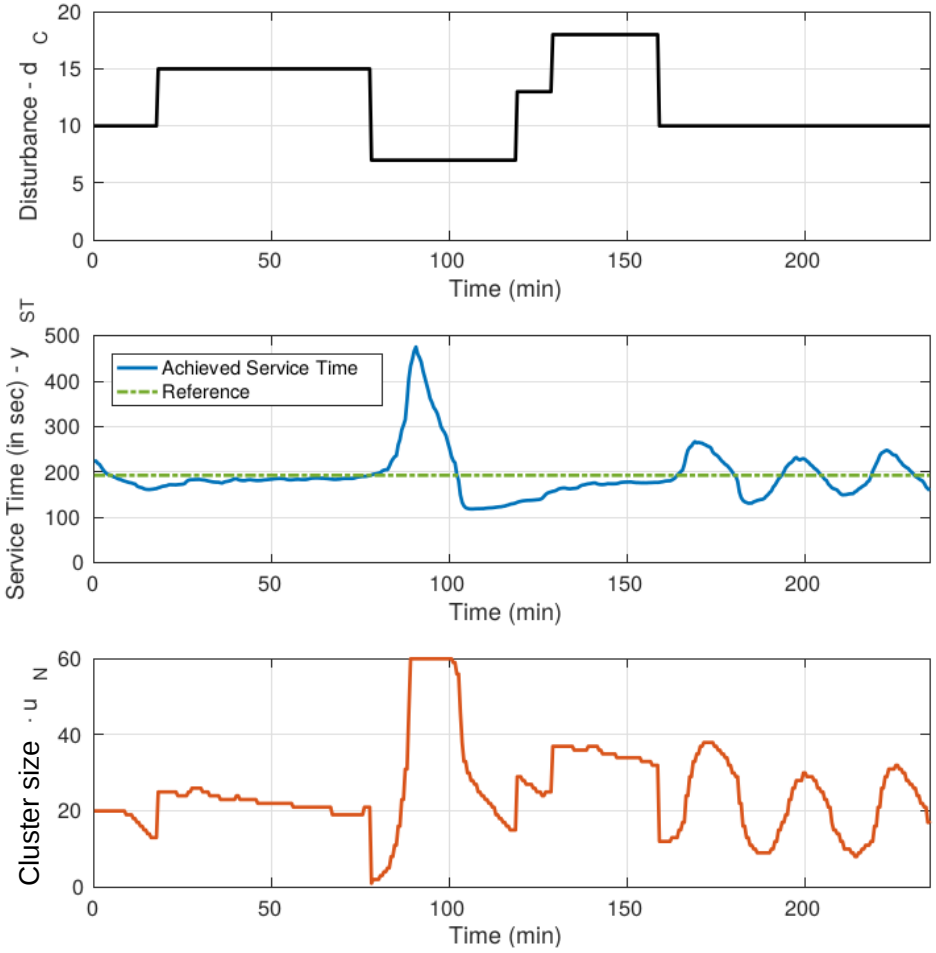
Formulation – Stability – Performance



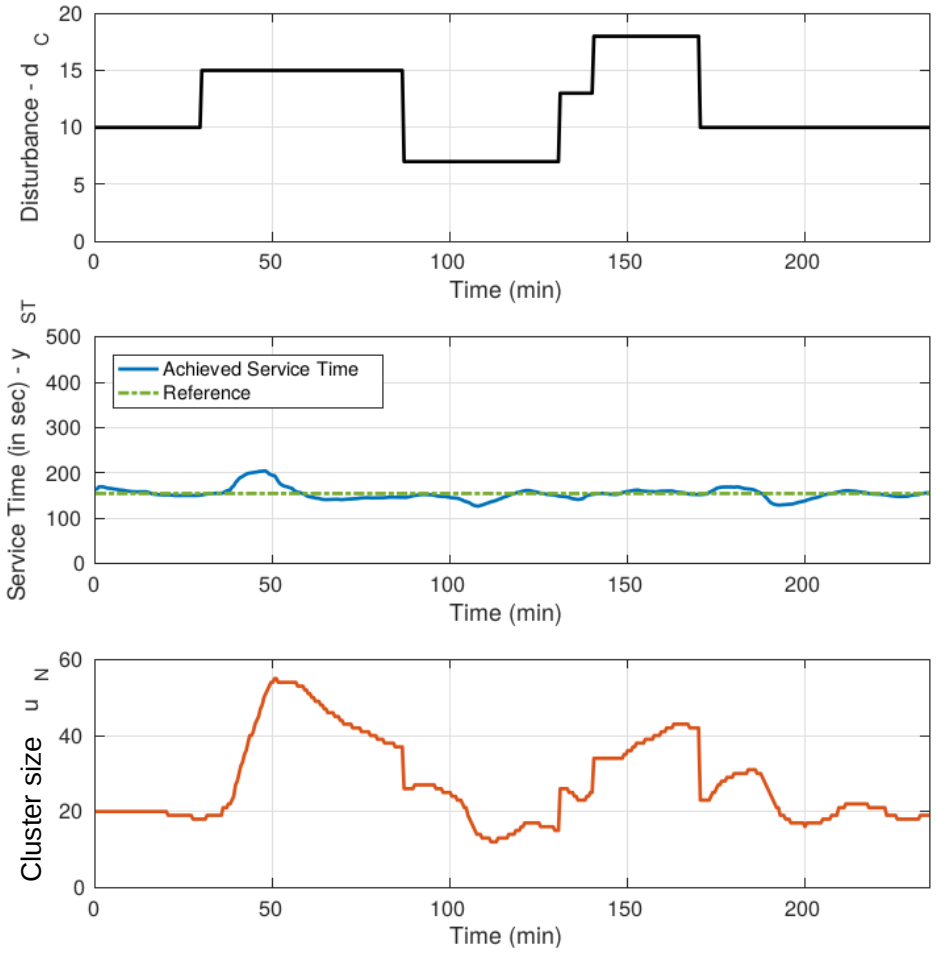
Robust Cloud Control

Formulation – Stability – Performance

- State of the Art [TCC, 2016]



- Adaptive Approach



Cost-aware Cloud Control

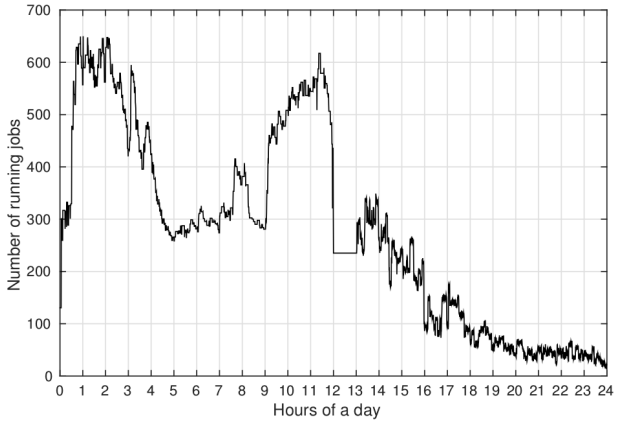
Formulation – Evaluation

- Objectives
  - Guarantee desired service time and availability
  - Reduce costs of cluster reconfiguration
- Formulation
  - Model Predictive Control
    - multi objective optimal control with constraints
  - Event-Based mechanism
    - reduce control reconfiguration
    - State of the art : control signal updates monitoring ? prediction ? properly deal with MIMO objectives ?
- Need for a new triggering mechanism
  - **Lyapunov Cost function** based Event Triggering Mechanism
    - Compute Lyapunov value if the control signal follows the previously computed trajectory
    - Compute Lyapunov value if the control signal is recomputed (= optimal)
    - Compare them (with a threshold)

Cost-aware Cloud Control

Formulation – Evaluation

- Stability
  - [CDC'16]
- Performance
  - Real 1-day workload (Ren, 2012)
  - Guaranteed service time and availaility
  - Events reduction
    - 90 % less cluster reconfiguration
  - Financial cost comparison Amazon pricing



Method	Fees	Extra costs compared to constrained cost based	
No control	5000$	3136$	**62.7%**
Time based (unconstrained)	1970$	107$	**5.4%**
Error based (unconstrained)	2020$	157$	**7.8%**
Cost based (unconstrained)	1867$	103$	**5.3%**
Constrained cost based	1863$	-	-

Outlines

I. Background

II. Location Privacy

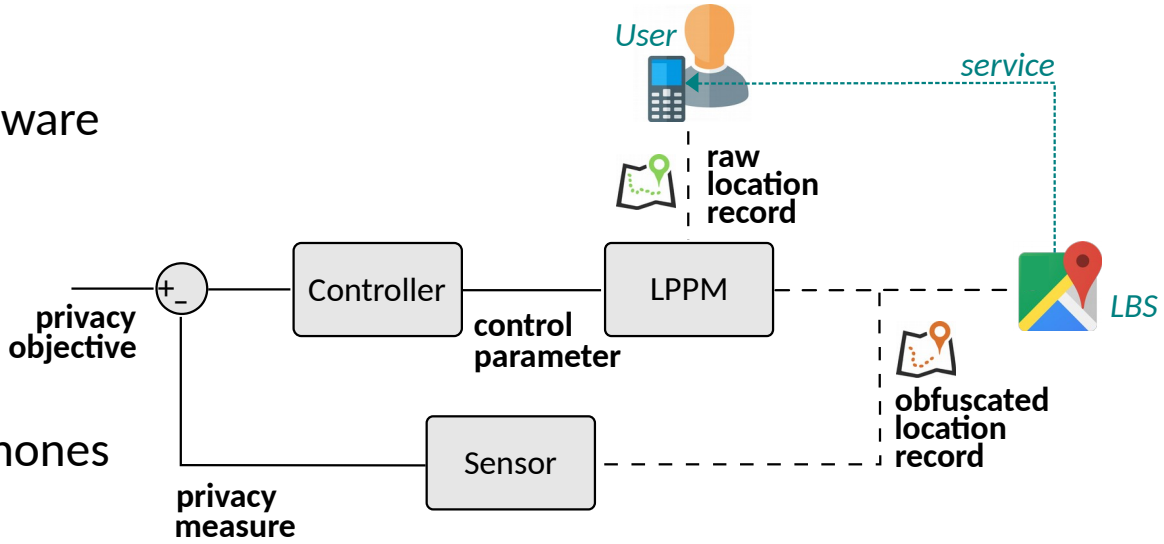
III. BigData Cloud Services

IV. Conclusions

- on Location Privacy Control
- on Cloud Services
- on Control of Computing Systems
- Perspectives

Location Privacy Control

- Contributions
  - Point-of Interest based Privacy
    - low level indicator of habits & identity
  - Automatic and robust privacy vs. utility trade-off
    - Database scenario: multi-objectives guarantees, evaluation over 770 users, faster than state of the art
    - Location broadcasting scenario: problem formulation, modeling, control proof of concept
- Perspectives
  - Predictive, optimal, utility-aware controller
  - Extended evaluation
  - Aditionnal utility and privacy notions
  - Implementation on smartphones

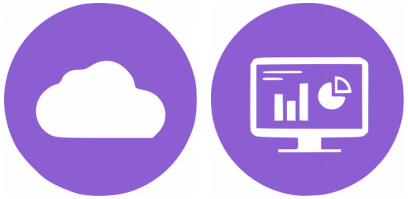


BigData Cloud Services

- Contributions
  - Resource provisionning and admission control
  - Adaptive service time monitoring for robustness against workload and environment changes
  - Multi-objectives optimal predictive controller with Lyapunov-based event triggering funtion
- Perspectives
  - Combination of robust and multi-objectives controllers
  - Addition of more metrics
    - 99th percentile of service time
  - Experimentations on a commercial cloud

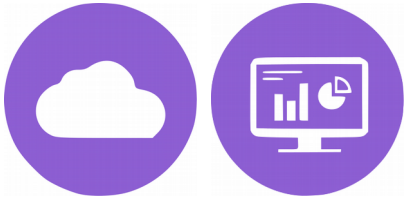
Control of Computing Systems

- Objectives
  - Automated and robust software adaptation with formal guarantees
- Contributions
  - Development of two use-cases
    - show a large variety of usage of control tool
    - contributions on complementary use-cases
  - State of the art challenges for the computing world
    - Complementary to other solutions
    - Control mathematical background enables formal guarantees
  - Theoretical contribution for the control community
    - New properties of systems require extended theory (scale, complexity, non-physical laws, etc.)



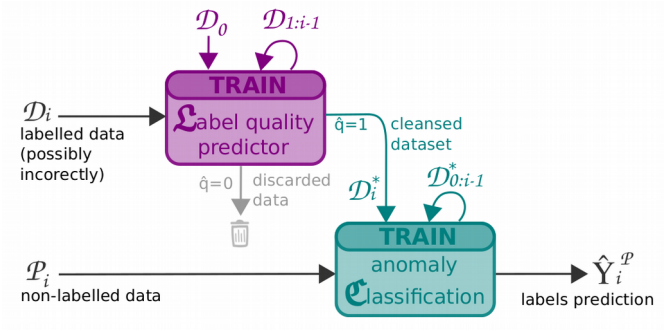
Lessons Learned

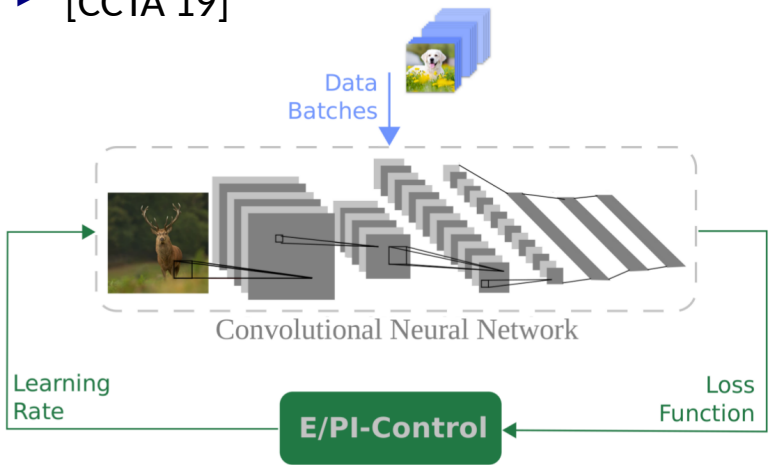
- Main Challenges
  - Problem Formulation
    - Inputs, outputs and cost functions are not intuitive
  - Modeling
    - Systems are not ruled by physics laws
  - Orders of magnitude of signals
  - Proved guarantees
- Limitations
  - Two distinct communities
    - vocabulary, communication venues, etc.
  - Dynamical system hypothesis



Perspectives

- Machine learning algorithms as a system to control
  - Internship @ IBM Research Zurich (July-Nov. 2018)
  - Objectives
    - Take the time-dynamics perspective
    - Explore more metrics than accuracy such as robusntess or privacy
    - Include the decision process in the modeling
  - Robust Learning on Unreliable Data
    - [NIPS'18] [DSN'19]
  - Feedback-based Training of Neural Networks
    - [CCTA'19]





Acknowledgment

- **Dr. Sonia Ben Mokhtar** (LIRIS lab, INSA-Lyon) on Location Privacy,
- **Dr. Mihaly Berekmeri** (Equifax UK) on modeling and control of Hadoop,
- **Dr. Robert Birke** (ABB Research) on data analytics and privacy,
- **Pr. Sara Bouchenak** (LIRIS lab, INSA-Lyon) on all the contributions of this manuscript,
- **Dr. Antoine Boutet** (Privatics, INRIA Lyon) on Location Privacy,
- **Dr. Lydia Y. Chen** (TU Delft) on data analytics and privacy,
- **Pr. Ioan D. Landau** (Gipsa-lab, Univ. Grenoble-Alpes) on adaptive control of Clouds,
- **Dr. Vincent Primault** (University College London) on Location Privacy,
- **Zilong Zhao** (Gipsa-lab, Univ. Grenoble-Alpes) on Machine Learning.

Publications

- International Journals
  - Submitted: IEEE Transaction on Automatic Control (**TAC**)
  - IEEE Transaction on Dependable and Secure Computing (**TDSC**) 2018
- International Conferences with Proceedings
  - 55th IEEE Conference on Decision and Control (**CDC** 2016), Las Vegas, United States
  - Conférence francophone d'informatique en parallélisme, architecture et système (**ComPAS** 2017), Sophia, France
  - **IFAC World Congress** 2017, Toulouse, France
  - 36th International Symposium on Reliable Distributed Systems (**SRDS** 2017), Hong Kong, SAR China
  - 18th IFIP International Conference on Distributed Applications and Interoperable Systems (**DAIS** 2018), Madrid, Spain
  - 2nd IEEE Conference on Control Technology and Applications (**CCTA** 2018), Copenhagen, Denmark
  - Continual Learning Workshop, Neural Information Processing Systems (**NIPS** 2018), Montréal, Canada
  - 49th IEEE/IFIP International Conference on Dependable Systems and Networks (**DSN** 2019), Portland, United States
  - 3rd IEEE Conference on Control Technology and Applications (**CCTA** 2019), Hong Kong, SAR China
- Posters & Abstracts
  - PhD Forum 34th International Symposium on Reliable Distributed Systems (**SRDS** 2015), Montreal, Canada
  - 11th International Workshop on Feedback Computing 2016, Wurzburg, Germany.
  - 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (**DSN** 2016), Toulouse, France
  - ACM/IFIP/USENIX **Middleware** conference 2016, Trente, Italy