SMART CITIES AS CYBER PHYSICAL SYSTEMS

C. G. Cassandras

Division of Systems Engineering, Dept. of ECE
Center for Information and Systems Engineering
Boston University
OUTLINE

- What is a “Smart City”?

- A Data-Driven Dynamic Resource Allocation Framework

- Examples of Smart City problems and solutions:
  - Adaptive Traffic Light Control
  - Smart Parking
  - Street Bump
  - Traffic control: eliminating the Price of Anarchy (PoA)
“SMART CITY” AS A CYBER-PHYSICAL SYSTEM

- Decision Making
- Data Collection
- Energy Management
- Safety
- Control and Optimization Actions
- Privacy
- Information Processing
- Decision Making
- Energy Management
- Sensor Networks
- Big Data

Christos G. Cassandras
CISE - CODES Lab. - Boston University
“SMART CITY” AS A CYBER-PHYSICAL SYSTEM

This is a HYBRID SYSTEM

Decision Making

Model

\[ \dot{x} = f(x,u,t) \]

Model

\[ x(t) \]

Christos G. Cassandras

CISE - CODES Lab. - Boston University
WHAT IS A “SMART CITY”?

“A city well performing in a forward-looking way in [economy, people, governance, mobility, environment, and living] built on the smart combination of endowments and activities of self-decisive, independent and aware citizens.”

Giffinger et al, 2007

Smart Sustainable Cities use information and communication technologies (ICT) to be more intelligent and efficient in the use of resources, resulting in cost and energy savings, improved service delivery and quality of life, and reduced environmental footprint—all supporting innovation and the low-carbon economy.

Cohen, 2014

Hitachi’s vision for the Smart Sustainable City seeks to achieve concern for the global environment and lifestyle safety and convenience through the coordination of infrastructure. Smart Sustainable Cities realized through the coordination of infrastructures consist of two infrastructure layers that support consumers’ lifestyles together with the urban management infrastructure that links these together using IT.

Hitachi Web, 2014

“We believe a city to be smart when investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance.”

Meijer and Bolívar, 2013
Smart Sustainable Cities use information and communication technologies (ICT) to be more intelligent and efficient in the use of resources, resulting in cost and energy savings, improved service delivery and quality of life, and reduced environmental footprint—all supporting innovation and the low-carbon economy.

*Cohen, 2014*
WHAT IS A “SMART CITY”?  

CREDIT: Fernando Livschitz

http://www.fastcodesign.com/3035870/filmmaker-creates-worlds-most-terrifying-traffic-intersection
WHAT IS REALLY “SMART” ?

COLLECTING DATA IS NOT “SMART”
- JUST A NECESSARY STEP TO BEING “SMART”

PROCESSING DATA TO MAKE GOOD DECISIONS IS “SMART”
WHAT IS A “SMART CITY”? 

• Ubiquitous \textit{wireless connectivity} of users and resources

• Abundance of \textit{real-time data} shared among users and resources

• Bring in \textit{feedback control mechanisms}

• Reduce/eliminate much of the \textit{infrastructure} (e.g., \textit{Connected Automated Vehicles})

• Achieve \textit{system-centric (social)} optimality rather than \textit{user-centric (selfish)} optimality - the \textit{SOCIAL} component in CPS

Some SMART things we can do
DATA-DRIVEN DYNAMIC RESOURCE ALLOCATION
DYNAMIC RESOURCE ALLOCATION

STATE \( x(t) \)

\[ x(t') = f[x(t), u(t), e(t)], \quad t' > t \]

CONTROL
- Reserve resource
- De-assign resource
- Change resource price
- Steer vehicle
- …

EVENTS
- New request
- Cancel request
- Resource freed
- …

\[
\begin{align*}
\min_{u(t)} & \quad J(x(t), u(t), t) \\
\text{s.t.} & \quad \text{user constraints} \\
& \quad \text{resource constraints}
\end{align*}
\]
DATA-DRIVEN STOCHASTIC OPTIMIZATION

CONTROL/DECISION (Parameterized by $\theta$)

$\theta_{n+1} = \theta_n + \eta_n \nabla L(\theta_n)$

$\nabla L(\theta)$

GRADIENT ESTIMATOR

$\nabla E[L(\theta)]$

REAL-TIME DATA

GOAL: $\max_{\theta \in \Theta} E[L(\theta)]$

PERFORMANCE $E[L(\theta)]$

DIFFICULTIES: - $E[L(\theta)]$ NOT available in closed form
- $\nabla L(\theta)$ not easy to evaluate
- $\nabla L(\theta)$ may not be a good estimate of $\nabla E[L(\theta)]$

Christos G. Cassandras

CISE - CODES Lab. - Boston University
DATA-DRIVEN STOCHASTIC OPTIMIZATION IN DES: INFINITESIMAL PERTURBATION ANALYSIS (IPA)

CONTROL/DECISION (Parameterized by $\theta$) \rightarrow \text{Discrete Event System (DES)} \rightarrow \text{PERFORMANCE $E[L(\theta)]$}

For many (but NOT all) DES:
- Unbiased estimators
- General noise distributions
- Simple on-line implementation

$\theta_{n+1} = \theta_n + \eta_n \nabla L(\theta_n)$

$\nabla L(\theta)$

Sample path

Model

Christos G. Cassandras

A general framework for an IPA theory in Hybrid Systems

\[ \theta_{n+1} = \theta_n + \eta_n \nabla L(\theta_n) \]
HYBRID AUTOMATA

STOCHASTIC HYBRID AUTOMATA

Event at time $\tau_k(\theta)$

Event at time $\tau_{k+1}(\theta)$

$k$th discrete state (mode)

$\dot{x} = f_k(x, \theta, t)$

$\theta$: control parameter, $\theta \in \Theta$ (system design parameter, parameter of an input process, or parameter that characterizes a control policy)
THE IPA CALCULUS

NOTATION:

\[ x'(t) = \frac{\partial x(\theta,t)}{\partial \theta}, \quad \tau'_k = \frac{d \tau_k(\theta)}{d \theta} \]

**IPA: THREE FUNDAMENTAL EQUATIONS**

System dynamics over \((\tau_k(\theta), \tau_{k+1}(\theta))\):
\[ \dot{x} = f_k(x, \theta, t) \]

1. \( x'(\tau_k^+) = x'(\tau_k^-) + [f_{k-1}(\tau_k^-) - f_k(\tau_k^+)] \cdot \tau_k' \)

2. \( x'(t) = e^{\int_{\tau_k}^t \frac{\partial f_k(u)}{\partial x} du} \left[ \int_{\tau_k}^t \frac{\partial f_k(v)}{\partial \theta} e^{-\int_{\tau_k}^v \frac{\partial f_k(u)}{\partial x} du} dv + x'(\tau_k^+) \right] \)

3. \( \tau_k' = 0 \quad \text{or} \quad \tau_k' = -\left[ \frac{\partial g}{\partial x} f_k(\tau_k^-) \right]^{-1} \left( \frac{\partial g}{\partial \theta} + \frac{\partial g}{\partial x} x'(\tau_k^-) \right) \)

Recall:
\[ x'(t) = \frac{\partial x(\theta, t)}{\partial \theta} \]
\[ \tau_k' = \frac{d \tau_k(\theta)}{d \theta} \]

\( g(x(\theta, \tau_k), \theta) = 0 \)

Switching function

*Some more complicated cases omitted*
IPA PROPERTIES

1. **ROBUSTNESS**
2. **DECOMPOSABILITY**
3. **SCALABILITY**

Yao and Cassandras, J. DEDS, 2011
IPA PROPERTIES

Back to performance metric:

\[ L(\theta) = \sum_{k=0}^{N} \int_{\tau_k}^{\tau_{k+1}} L_k(x, \theta, t) dt \]

NOTATION:

\[ L'_k(x, \theta, t) = \frac{\partial L_k(x, \theta, t)}{\partial \theta} \]

Then:

\[ \frac{dL(\theta)}{d\theta} = \sum_{k=0}^{N} \left[ \tau'_{k+1} \cdot L_k(\tau_{k+1}) - \tau'_k \cdot L_k(\tau_k) + \int_{\tau_k}^{\tau_{k+1}} L'_k(x, \theta, t) dt \right] \]

What happens at event times

What happens between event times
THEOREM 1: If either 1,2 holds, then \( \frac{dL(\theta)}{d\theta} \) depends only on information available at event times \( \tau_k \):

1. \( L(x, \theta, t) \) is independent of \( t \) over \([\tau_k(\theta), \tau_{k+1}(\theta)]\) for all \( k \)

2. \( L(x, \theta, t) \) is only a function of \( x \) and for all \( t \) over \([\tau_k(\theta), \tau_{k+1}(\theta)]\):

\[
\frac{d}{dt} \frac{\partial L_k}{\partial x} = \frac{d}{dt} \frac{\partial f_k}{\partial x} = \frac{d}{dt} \frac{\partial f_k}{\partial \theta} = 0
\]

IMPLICATION: - Performance sensitivities can be obtained from information limited to event times, which is easily observed

- No need to track system in between events!
1. ROBUSTNESS

EVENTS

OBVIOUS: Evaluating $x(t; \theta)$ requires full knowledge of $w$ and $f$

NOT OBVIOUS: $\frac{dx(t; \theta)}{d \theta}$ may be independent of $w$ and $f$

It often depends only on:
- event times $\tau_k$
- possibly $f(\tau_{k+1})$

$\ddot{x} = f(x, u, w, t; \theta)$
THEOREM 2: Suppose an endogenous event occurs at $\tau_k$ with switching function $g(x, \theta)$.

If $f_k(\tau_k^+) = 0$, then $x'(\tau_k^+)$ is independent of $f_{k-1}$.

If, in addition, $\frac{dg}{d\theta} = 0$ then $x'(\tau_k^+) = 0$.

IMPLICATION: Performance sensitivities are often reset to 0, $\Rightarrow$ sample path can be conveniently decomposed.
3. SCALABILITY

OBSERVATION: IPA is entirely *event-driven*

⇒ scales with event set size, not state space!

1. \[ x'(\tau_k^+) = x'(\tau_k^-) + [f_{k-1}(\tau_k^-) - f_k(\tau_k^+)] \cdot \tau_k' \]

2. \[ x'(\tau_{k+1}^-) = e^{\int_{\tau_k}^{\tau_{k+1}} \frac{\partial f_k(u)}{\partial x} du} \left[ \int_{\tau_k}^{\tau_{k+1}} \frac{\partial f_k(v)}{\partial \theta} dv \right] x'(\tau_k^-) + x'(\tau_k^+) \]

3. \[ \tau_k' = 0 \quad \text{or} \quad \tau_k' = \left[ \frac{\partial g}{\partial x} f_k(\tau_k^-) \right]^{-1} \left( \frac{\partial g}{\partial \theta} + \frac{\partial g}{\partial x} x'(\tau_k^-) \right) \]
In many cases:

- **No need for a detailed model** (captured by $f_k$) to describe state behavior in between events

- This explains why *simple abstractions of a complex stochastic system* can be adequate to perform sensitivity analysis and optimization, as long as event times are accurately observed and local system behavior at these event times can also be measured
A SMART CITY APPLICATION: ADAPTIVE TRAFFIC LIGHT CONTROL
A basic binary switching control (GREEN – RED) problem with a long history...

- Mixed Integer Linear Programming (MILP) [Dujardin et al, 2011]
- Extended Linear Complementarity Problem (ELCP) [DeSchutter, 1999]
- MDP and Reinforcement Learning [Yu et al., 2006]
- Game Theory [Alvarez et al., 2010]
- Evolutionary algorithms [Taale et al., 1998]
- Fuzzy Logic [Murat et al., 2005]
- Expert Systems [Findler and Stapp, 1992]

- Perturbation Analysis
• **Perturbation Analysis** [Panayiotou et al., 2005]
  [Geng and Cassandras, 2012]

Use a Hybrid System Model: **Stochastic Flow Model (SFM)**

Aggregate states into *modes* and keep only events causing mode transitions
Traffic light control:
\[ \theta = [\theta_1, \theta_2, \theta_3, \theta_4] \]

**GREEN** light cycle at queue \( n = 1, 2, 3, 4 \)

**OBJECTIVE:**
Determine \( \theta \) to minimize total weighted vehicle queues

\[
\min_{\theta} J_T(\theta) = \frac{1}{T} E \left[ \sum_{n=1}^{4} \int_{0}^{T} w_n x_n(\theta, t) dt \right]
\]
\[
\min_{\theta} J_T(\theta) = \frac{1}{T} E \left[ \sum_{n=1}^{4} \int_0^T \omega_n x_n(\theta, t) \, dt \right] = \frac{1}{T} E[L_T(\theta)]
\]

IPA APPROACH:

- Observe events and event times, estimate \( \frac{dJ_T(\theta)}{d\theta} \) through \( \frac{dL_T(\theta)}{d\theta} \)

- Then, \( \theta_{n+1} = \theta_n + \eta_n \frac{dL_T(\theta_n)}{d\theta} \)
\[ \dot{z}_n(t) = \begin{cases} 1 & \text{if } 0 < z_n(t) < \theta_n \text{ or } z_{\bar{n}}(t) = \theta_{\bar{n}} \\ 0 & \text{otherwise} \end{cases} \]

\[ z_n(t^+) = 0 \text{ if } z_n(t) = \theta_n \]
### HYBRID SYSTEM STATE DYNAMICS

**[RESOURCE DYNAMICS]**

\[
\dot{z}_n(t) = \begin{cases} 
1 & \text{if } 0 < z_n(t) < \theta_n \text{ or } z_n(t) = \theta_n \\
0 & \text{otherwise}
\end{cases}
\]

\[
z_n(t^+) = 0 \text{ if } z_n(t) = \theta_n
\]

**Define:**

\[
G_n(t) = \begin{cases} 
1 & \text{if } 0 < z_n(t) < \theta_n \text{ or } z_n(t) = \theta_n \\
0 & \text{otherwise}
\end{cases}
\]

**[USER DYNAMICS]**

\[
\dot{x}_n(t) = \begin{cases} 
\alpha_n(t) & \\
0 & \text{if } G_n(t) = 1
\end{cases}
\]

**Vehicle departure rate process**

\[
\alpha_n(t) \quad \beta_n(t)
\]

**Vehicle arrival rate process**

**IPA ROBUSTNESS:**

\[
\alpha_n(t), \beta_n(t) \text{ do not have to be known!}
\]

GREEN: queue \(n\)
TYPICAL SIMULATION RESULTS

9-fold cost reduction

Traffic pattern changes

Adaptivity
EXTENSIONS

- Two intersections with blocking
  [Geng and Cassandras, J. DEDS, 2015]

- Quasi-Dynamic TLC: assume partially observable queues
  [Fleck, Cassandras and Geng, IEEE TCST, 2016]

- Network of intersections: exploit IPA SCALABILITY property
• Automatically adapt RED/GREEN light cycles based on observed data
• Predict and alleviate congestion over entire urban network
• Reduce waiting times, congestion
• Reduce pollution and fuel waste
Two ways of looking at this problem:

1. Control Traffic Lights (infrastructure intensive)
2. Control speed/acceleration of vehicle assuming connectivity between vehicles (V2V) and traffic lights (V2I)
   (e.g., adjust speed to make a GREEN just in time)
SMART PARKING

iPhone app
30% of vehicles on the road in the downtowns of major cities are cruising for a parking spot. It takes the average driver 7.8 minutes to find a parking spot in the downtown core of a major city.


GUIDANCE-BASED PARKING – DRAWBACKS...

**Drivers:**
- May not find a vacant space
- May miss better space
- Processing info while driving

**City:**
- Imbalanced parking utilization
- May create ADDED CONGESTION (as multiple drivers converge to where a space exists)

 Searching for parking  ⇒  Competing for parking
Find best parking spot for DESTINATION A

BEST PARKING SPOT

LEAST DISTANCE from A +
LEAST COST +
RESERVE IT

[Christos G. Cassandras, CISE - CODES Lab. - Boston University]

[Geng and Cassandras, IEEE Trans. on Intelligent Transportation Systems, 2013]
SMART PARKING – IMPLEMENTATION

- Parking space availability detection
  - Standard sensors (e.g., magnetic, cameras)
  - Wireless sensor networking

- Vehicle localization
  - GPS

- System-Driver communication
  - Smartphone
  - Vehicle navigation system

- Parking reservation
  - Red/Green/Yellow light system
Currently in operation at BU garage
(with Smartphone app: BU Smart Parking)


2011 IBM/IEEE Smarter Planet Challenge prize

Currently in operation at BU garage
(with Smartphone app: BU Smart Parking)
STREET BUMP: DETECTING “BUMPS” THROUGH SMARTPHONES + DATA ANALYTICS

2014 IBM/IEEE Smarter Planet Challenge prize

iPhone app
STREET BUMP – PROCESSING “BIG DATA”

• Detect obstacles using iPhone accelerometer and GPS ⇒ no infrastructure needed

• Send to central server through Street Bump app

• Process data to classify obstacles:
  Anomaly detection and clustering algorithms, similar to cybersecurity problems

• Detect “actionable” obstacles

• Prioritize and dispatch Smart City crews to fix problems:

  DATA-DRIVEN DYNAMIC RESOURCE ALLOCATION

[Brisimi et al, IEEE Access, 2016]
Methodologies used:

- Anomaly detection, Machine Learning algorithms
- Bump signal signature analysis: REGULARITY METRIC
- Bump signal randomness content: ENTROPY METRIC

NON-ACTIONABLE
(Flat Casting)

ACTIONABLE
(Pothole)
$AI = 0.5MSE + 0.5H(x)$

- $\lambda = 0.5$
- Truly actionable (T): 88/100 (88%)
- False Alarm (F): 12%

Normalized Comb. of MSE & Entropy - $\lambda = 0.5$
SHARING RESOURCES: THE “PRICE OF ANARCHY”
TRAFFIC CONTROL

100-km Chinese traffic jam enters Day 9

The BU Bridge mess, Boston, MA (simulation using VISSIM)
WHY CAN’T WE IMPROVE TRAFFIC…

... EVEN IF WE KNOW THE ACHIEVABLE OPTIMUM IN A TRAFFIC NETWORK ???

Because:

• **Not enough controls** (traffic lights, tolls, speed fines)
  → No chance to unleash the power of feedback!

• **Not knowing other drivers’ behavior** leads to poor decisions
  (a simple game-theoretic fact)
  → Drivers seek individual *(selfish)* optimum,
    not system-wide *(social)* optimum

---

Christos G. Cassandras
GAME-CHANGING OPPORTUNITY:
CONNECTED AUTOMATED VEHICLES (CAVs)

NO TRAFFIC LIGHTS, NEVER STOP...

FROM (SELFISH) “DRIVER OPTIMAL”
TO (SOCIAL) “SYSTEM OPTIMAL”
TRAFFIC CONTROL
HOW TO MEASURE THE PoA?

Under USER-CENTRIC control, $x^\text{user}_a$ is the equilibrium flow.

Under SYSTEM-CENTRIC control, $x^\text{social}_a$ is the equilibrium flow.
HOW TO MEASURE THE PoA?

\[
\text{PoA} = \frac{\sum_{\text{all } a} x_a^{\text{user}} t(x_a^{\text{user}})}{\sum_{\text{all } a} x_a^{\text{social}} t(x_a^{\text{social}})} \geq 1
\]

Can we measure/estimate the PoA?
DIFFICULTIES AS THINGS NOW STAND…

- We don’t know user COST FUNCTIONS
- We don’t know user ORIGIN-DESTINATION pairs (No DEMAND model)

- We can’t solve the SYSTEM OPTIMALITY problem

We can’t exploit CAVs

We can’t assess the value of investing in CAV-based technologies, since we can’t evaluate the PRICE OF ANARCHY

BUT WE DO HAVE PLENTY OF DATA….
INVERSE OPTIMIZATION PROBLEM

KEY IDEA:

BACKWARD optimization:
- Data reveal a (selfish) equilibrium (Wardrop/Nash equilibrium)
- What are the (virtual) cost functions that best fit the data and lead to this equilibrium?

FORWARD optimization:
- Use these cost functions to find (social) optimal traffic flows
- Estimate the PRICE OF ANARCHY
INTERPRETATION:
We can improve traffic by more than 100% if we can direct vehicles (e.g., using CAVs)

Zhang et al, IEEE CDC, 2016 - MoA22
A DECENTRALIZED OPTIMAL CONTROL FRAMEWORK FOR CAVs

A story for another time, but here is what the end of its first chapter looks like...

[Zhang et al, ACC, 2016]
WHO NEEDS TRAFFIC LIGHTS?

With traffic lights

With decentralized control of CAVs

One of the worst-designed double intersections ever...
(BU Bridge – Commonwealth Ave, Boston)
WHO NEEDS TRAFFIC LIGHTS?

How about pedestrians?
“Smart Cities” are complex CYBER-PHYSICAL systems that can be studied in a stochastic hybrid system setting.

Capitalize on WIRELESS NETWORKING + BIG DATA + DATA-DRIVEN CONTROL and OPTIMIZATION METHODS.

“CONNECTED VEHICLES” provide a tremendous opportunity for feedback methods, game theoretic approaches, no infrastructure.

What about HUMANS? Need to expand to CYBER - SOCIAL - PHYSICAL systems.
STUDENTS:

…and COLLEAGUES:
Y. Wardi, I. Paschalidis, A. Malikopoulos, R. Su, Q.S. Jia