Learning and Control

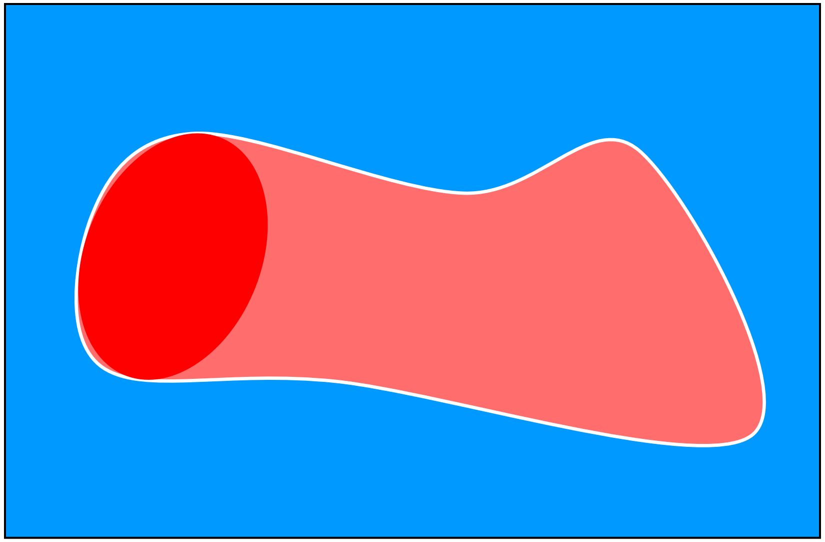
Anil Aswani, Haomiao Huang, Zhengyuan Zhou, Claire Tomlin, Shankar Sastry

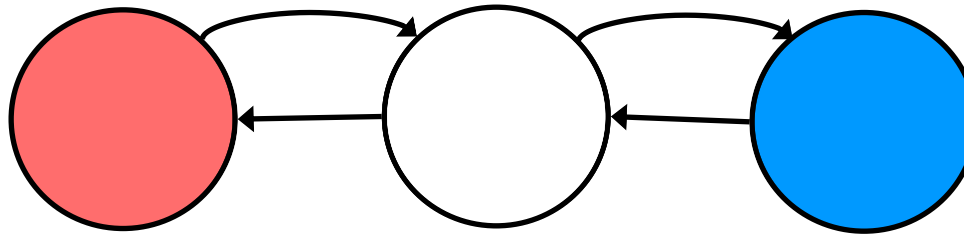
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley

April 12, 2013

Results presented here have been supported by:
NSF CPS ActionWebs, ONR, AFOSR

System verification and control design





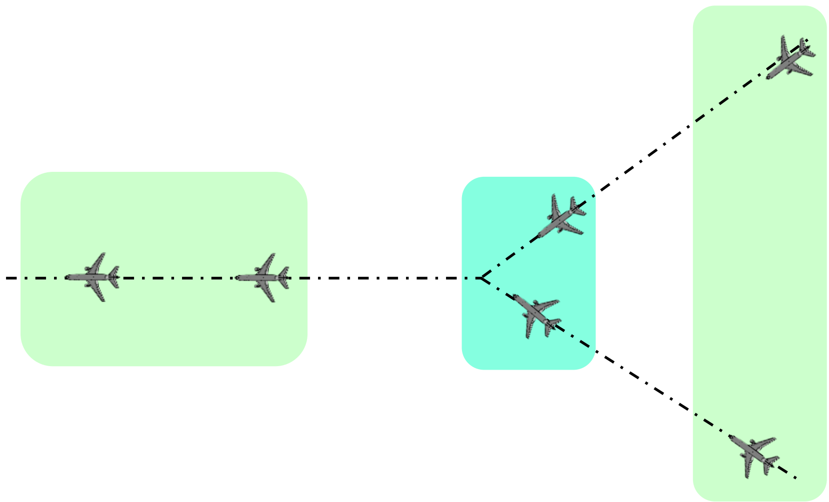
In red, system may become unsafe

In blue, system will stay safe

- much of control theory is concerned with certifying system behavior (stability, robustness, reachability…)
- usually performed on simple, abstract models
- tools preclude complex dynamics, such as human behavior, economic implications

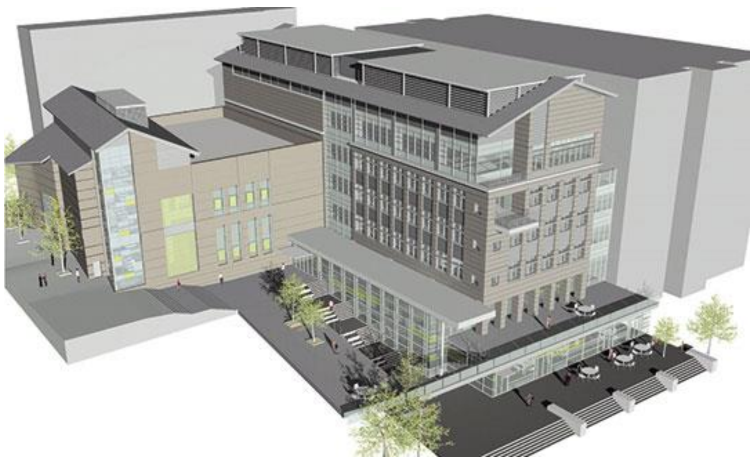
Examples

Air traffic control



- Controllers, pilots, passengers
- Grouping and conflict classification
- Uncertainty

Energy Efficient Buildings

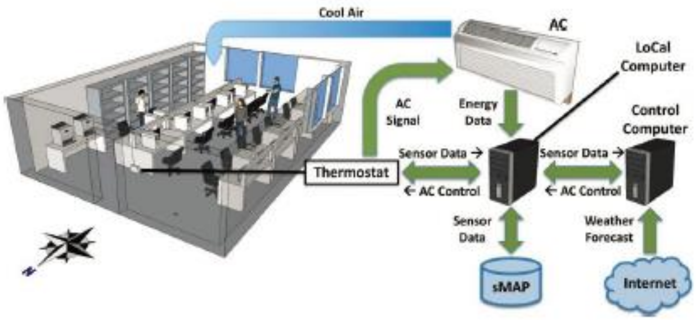


- Controller, occupants
- Interlinked spaces with time-varying objectives
- Uncertainty

[with Hamsa Balakrishnan]

Uncertainty from Humans

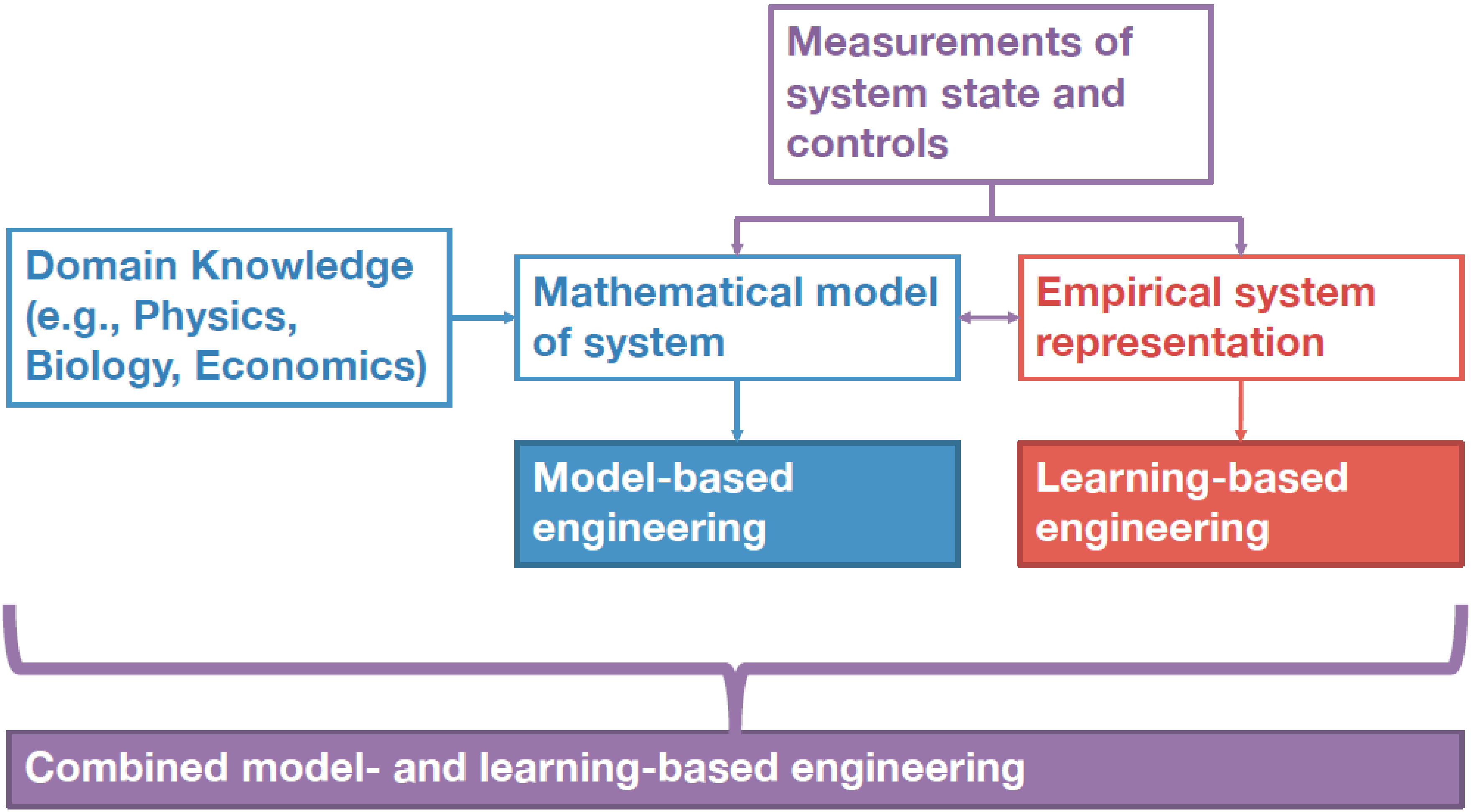
- Interplay between
  - Physics
  - Humans
  - Computation
- Tackling uncertainty will
  - Improve efficiency
  - Increase performance
  - Enhance robustness



Smart Buildings
(Aswani, et al., Proc. IEEE, 2011)

High-performance Systems
(Aswani, et al., ACC, 2012)

Research Philosophy



Learn models from data…

… but satisfy control requirements while learning

- **Control requirements:**
  - A nominal model with error bounds
  - Reachable sets computed to ensure properties hold in worst case
  - Reachable sets computed using Model Predictive Control (MPC)
- **Performance:**
  - Use online learning to update model
  - Cost function used to generate control action within the safe set
- Insight: performance and constraint satisfaction can be decoupled using reachability analysis
- **Learning-based (LB) Model Predictive Control (MPC)**

Constraint satisfaction through Reachability

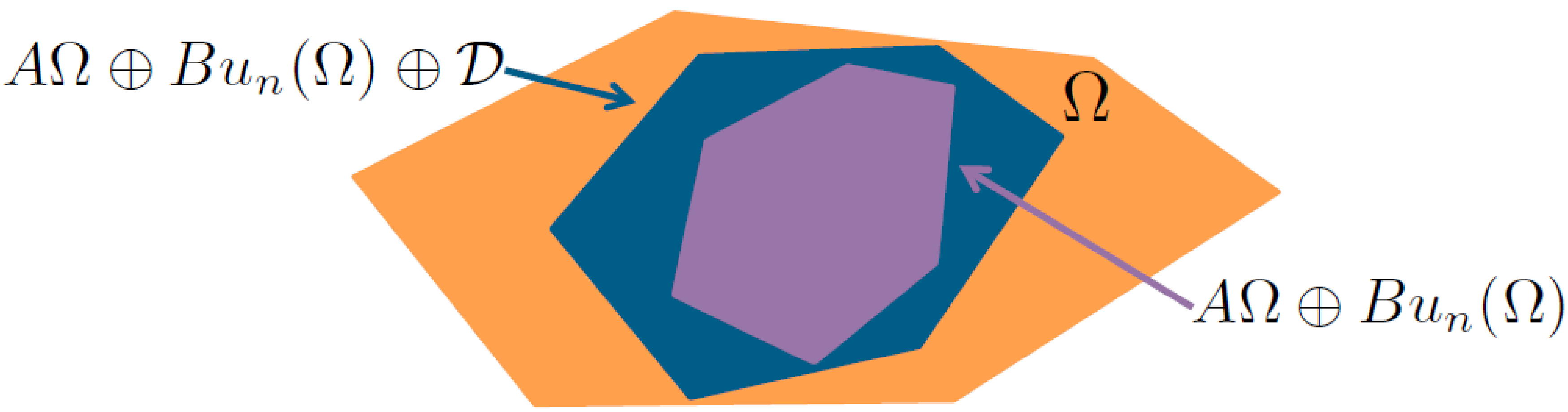
- **Reachable set Ω**
  - **Constraint satisfaction**

$$x_n \in \Omega \subseteq \mathcal{X}, u_n(x_n) \in \mathcal{U}, \forall n \geq 0$$

  - **Disturbance invariance**

$\forall x_n \in \Omega \; \exists u_n(x_n) \in \mathcal{U} :$

$$x_{n+1} = Ax_n + Bu_n + d_n \subseteq \Omega, \forall d_n \in \mathcal{D}$$



Learning-based Model Predictive Control

- Unknown system dynamics represented using an oracle $\mathcal{O}$
- At each time step
  - Optimization solved, Oracle updated

LBMPC	Performance	
$u_m^* = \arg\min$	$J(\tilde{x}_{m+1}, \ldots, \tilde{x}_{m+N}, u_m, \ldots, u_{m+N-1})$	• Cost function
s.t.	$\tilde{x}_{n+1} = A\tilde{x}_n + Bu_n + \mathcal{O}_m(\tilde{x}_n, u_n)$	• Learned model
	$x_{m+k} \in \mathcal{X} \ominus \mathcal{R}_i; x_{m+N} \in \mathcal{X}_N \ominus \mathcal{R}_N$	• Constraints and uncertainty
	$u_{m+k} = Kx_{m+k} + c_{m+k} \in \mathcal{U} \ominus K\mathcal{R}_i$	
	$x_{n+1} = Ax_n + Bu_n$	• Original model
	Safety	

[Aswani et al., Automatica 2013]

Learning-based Model Predictive Control

Theorem

If the reach set can be computed, then the closed-loop system that uses the control input computed by LBMPC satisfies all state and input constraints and is stable.

Theorem

If the oracle is bounded and continuous, then the value function of the LBMPC optimization problem is continuous.

Theorem

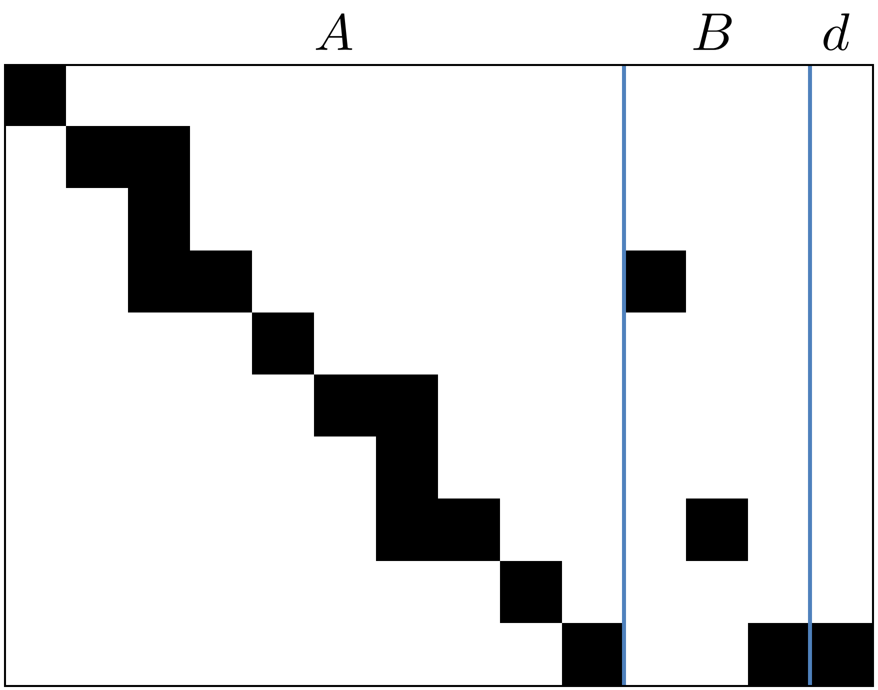
If the reach set can be computed, the linear MPC with the nominal model is asymptotically stable, and the oracle is bounded and continuous; then the closed-loop system is robustly asymptotically stable (RAS) stable with respect to the modeling error and disturbance.

[Aswani et al., Automatica 2013]

Example 1: Learning to fly

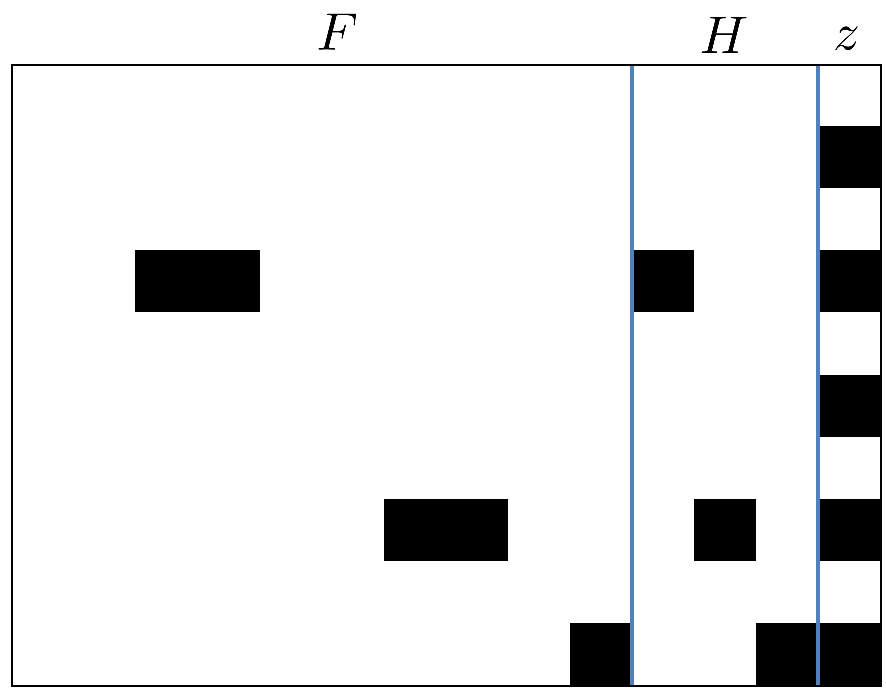
- Linear model

$$x_{n+1} = Ax_n + Bu_n + d$$



- Statistics augments physics

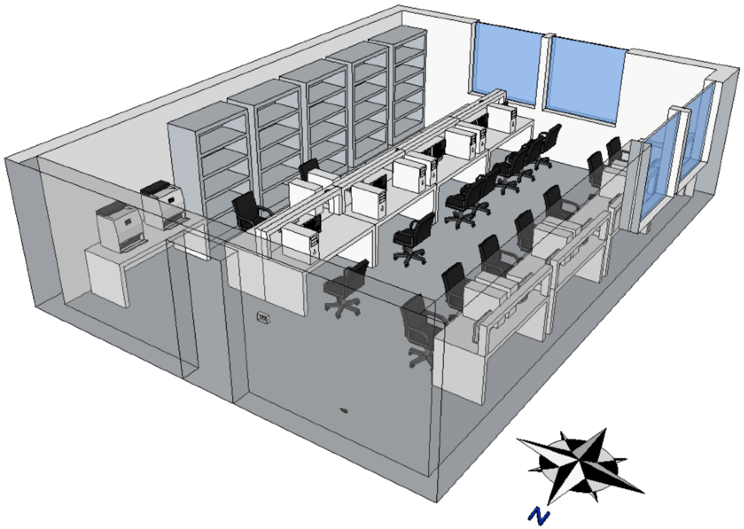
$$\mathcal{O}_m = Fx_n + Hu_n + z$$



Example 1: Learning to fly

video

Example 2: Energy-efficient buildings



- 640 sq. ft. computer space
- Networked thermostat
- Newton's law of cooling with heating load

- 141,000 sq. ft. building
- Seven floors of mixed-use space
- Convective cooling with heating load

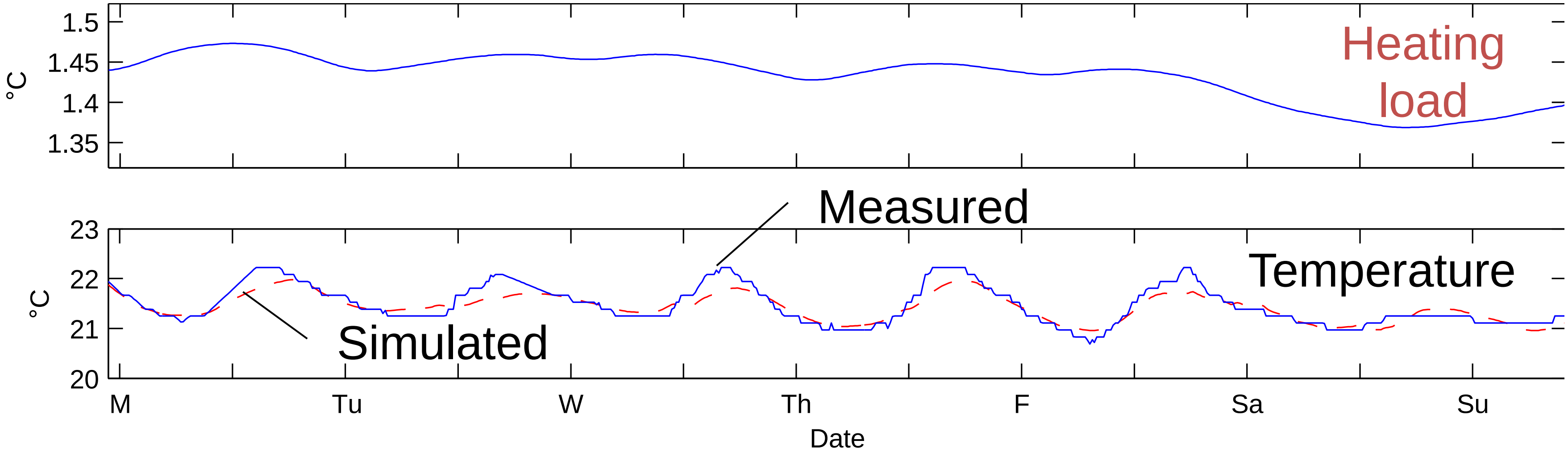
Berkeley Retrofitted and Inexpensive HVAC Testbed for Energy Efficiency (BRITE) [with Aswani, Culler, Taneja, Krioukov]

Temperature Modeling

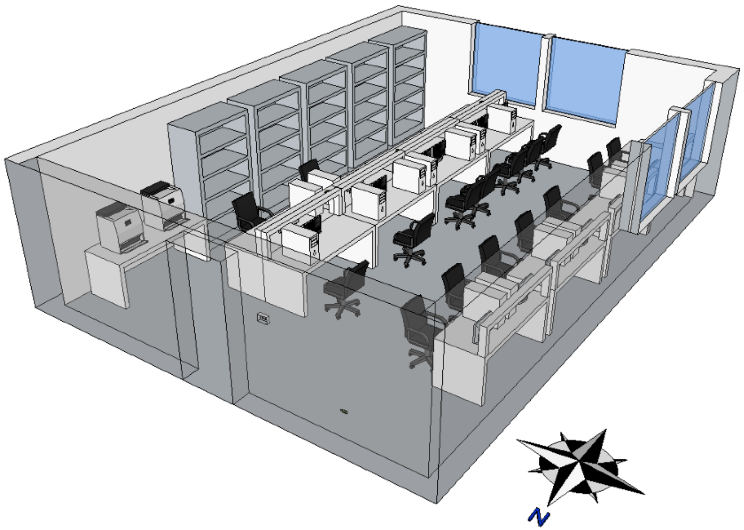
- Semi-parametric regression modeling
  - Parametric: Newton's law of cooling
  - Nonparametric: Heating load

$$T_{n+1} = AT_n + B_1 u_n + B_2 w_n + q_n$$

  - Estimate heating load using only temperature measurements of thermostat



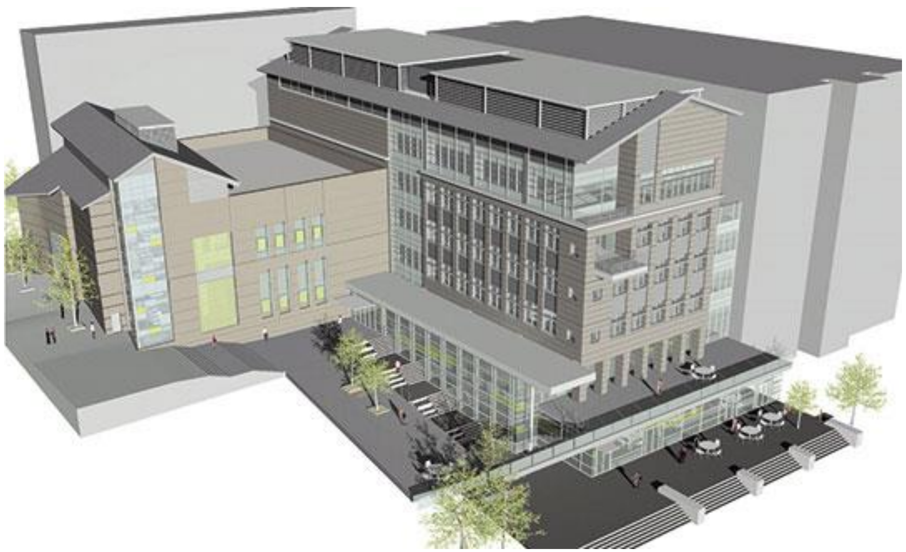
Experiments on BRITE

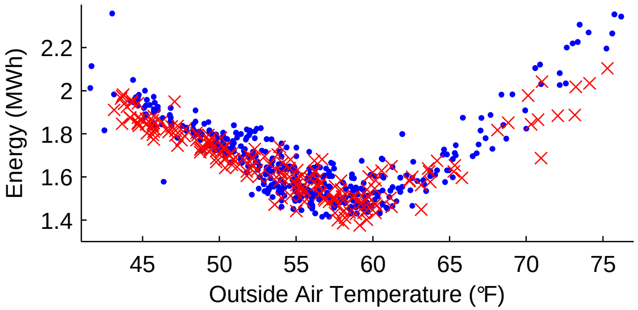


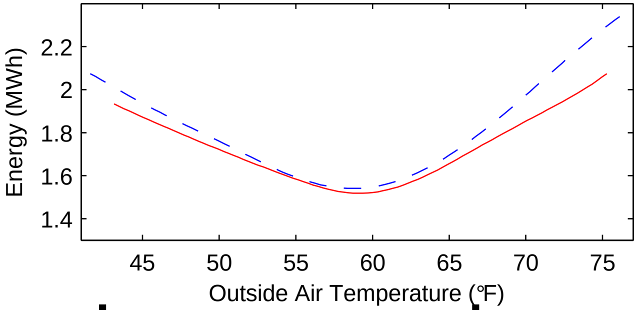
Experiment	Method	Energy
Thermostat Controller	LBMPC	23.6 kWh
	Thermostat	32.6 kWh
LBMPC Controller	LBMPC	11.8 kWh
	Thermostat	34.5 kWh

- LBMPC provides significant energy savings
- Simulations and experiments used to compare controllers

Experiments on BRITE-S





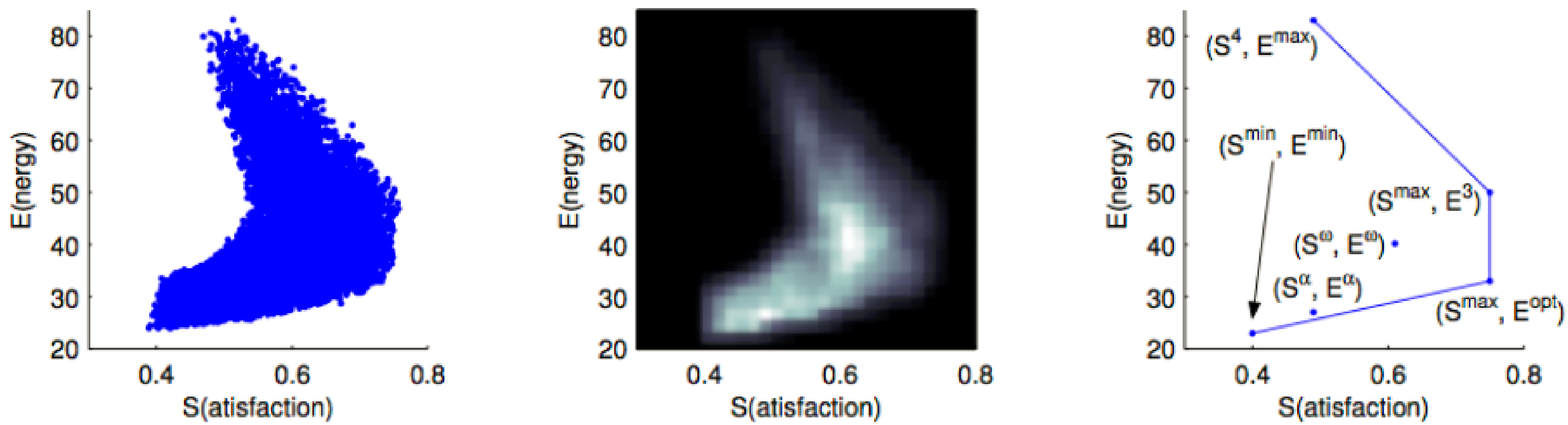


- Comparisons using many experiments and nonparametric hypothesis testing
- 1.5MWh energy saved on average per day
  - Statistically significant (p=0.002)
  - 95% Confidence Interval of 1MWh to 2MWh savings

Incentive Design for Efficient Quality of Service

Key Question

How can incentive schemes that encourage energy-efficient operation of buildings without degradations in quality of service be designed, while also preventing gaming (e.g., moral hazards)?

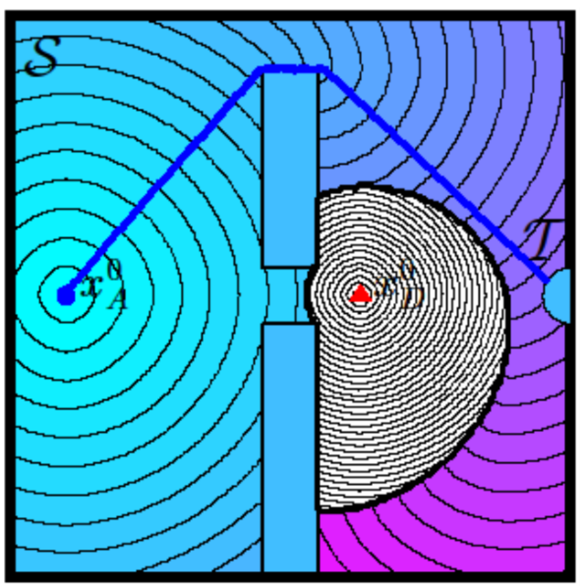


(a) Scatter Plot (b) Density Plot (c) Key Points

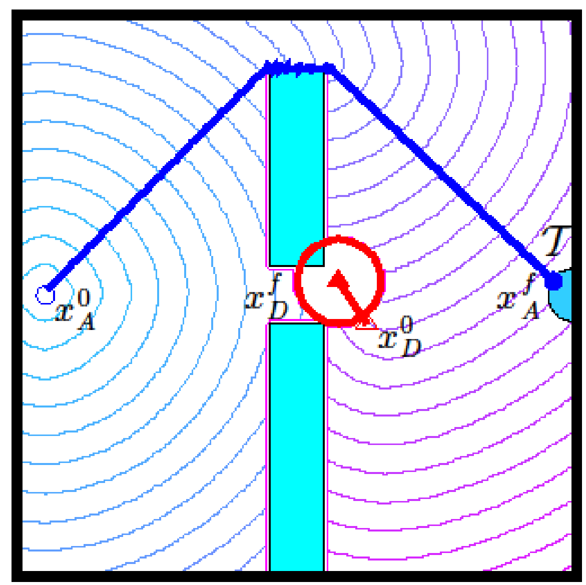
[Aswani and Tomlin, Allerton, 2012]

Fast Certification

- Reachability specifications typically exponential in dimension of continuous state space
- Fast approximations of “upper value” and “lower value” using an open loop game: $\underline{\underline{v}}(\mathbf{x^0}) \leq \underline{v}(\mathbf{x^0}) \leq V(\mathbf{x^0}) \leq \overline{v}(\mathbf{x^0})$
  - **Result 1:** Open-Loop upper value can be computed exactly [Zhou, Huang, Ding, Takei, Tomlin, submitted 2012]
  - **Result 2:** A lower bound for the open-loop lower value can be computed [Zhou, Takei, Huang, Tomlin, CDC 2012]



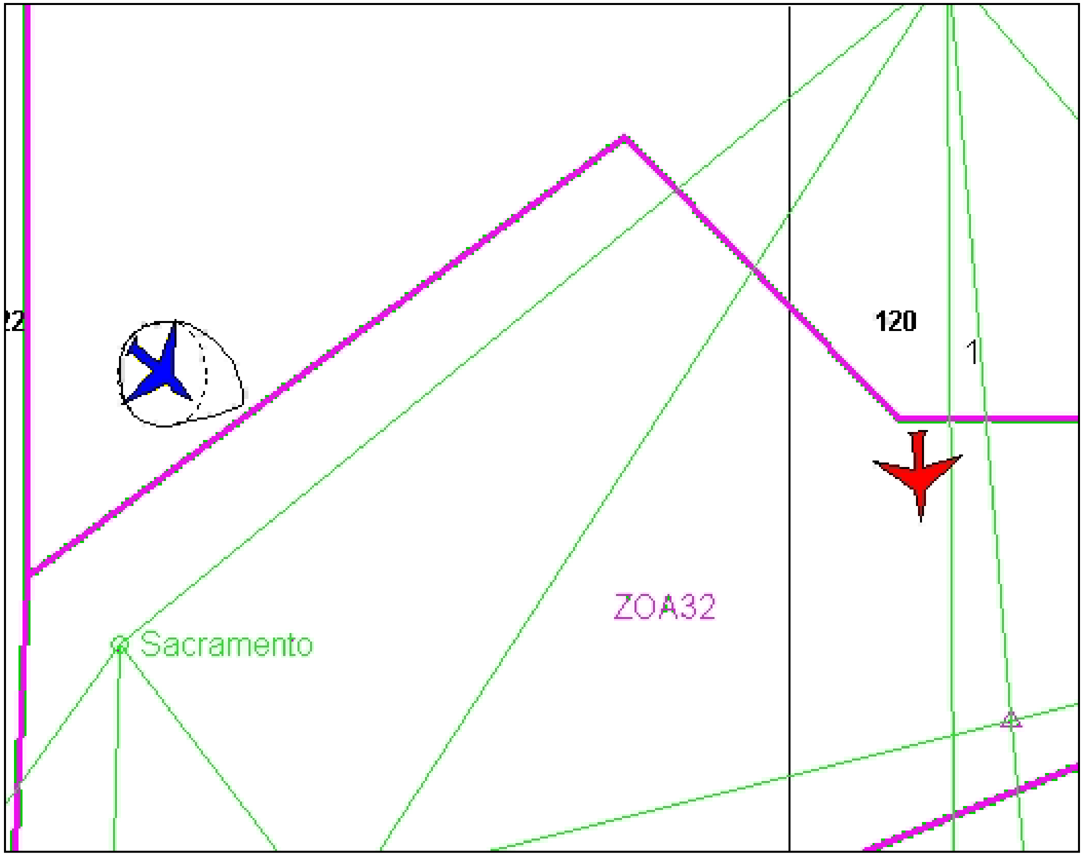
(upper)



(lower)

Characterizing human interaction

Example: reachability in air traffic control



[from Alex Bayen, a long time ago]

Example 3: Automation to aid controllers

- Infeasible to get data from real controllers
- Most experiments use retired controllers or student volunteers
- Retired controllers are rare, students get bored, where to get more data?

Contrails: Air traffic control game for Android

Replay Engine on Server

Trajectories, aircraft states, player inputs

[Huang and Tomlin 2012]

A Typical ATC experiment[1]

28 participants
168 trials (6 each)

Local US college students

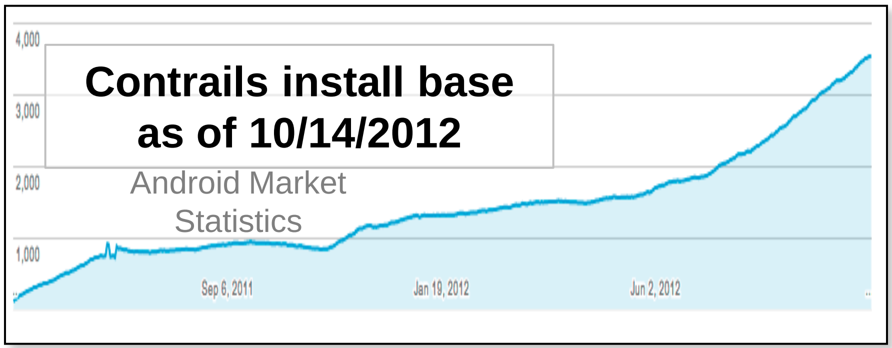
Max individual sample (est): **100** planes

Contrails to date

3544 active installs
63,583 games played

10+ countries

Most active user: **9489** planes



Contrails install base as of 10/14/2012

Users by country, as of 10/14/2012

Leaderboard

Top controllers from around the world

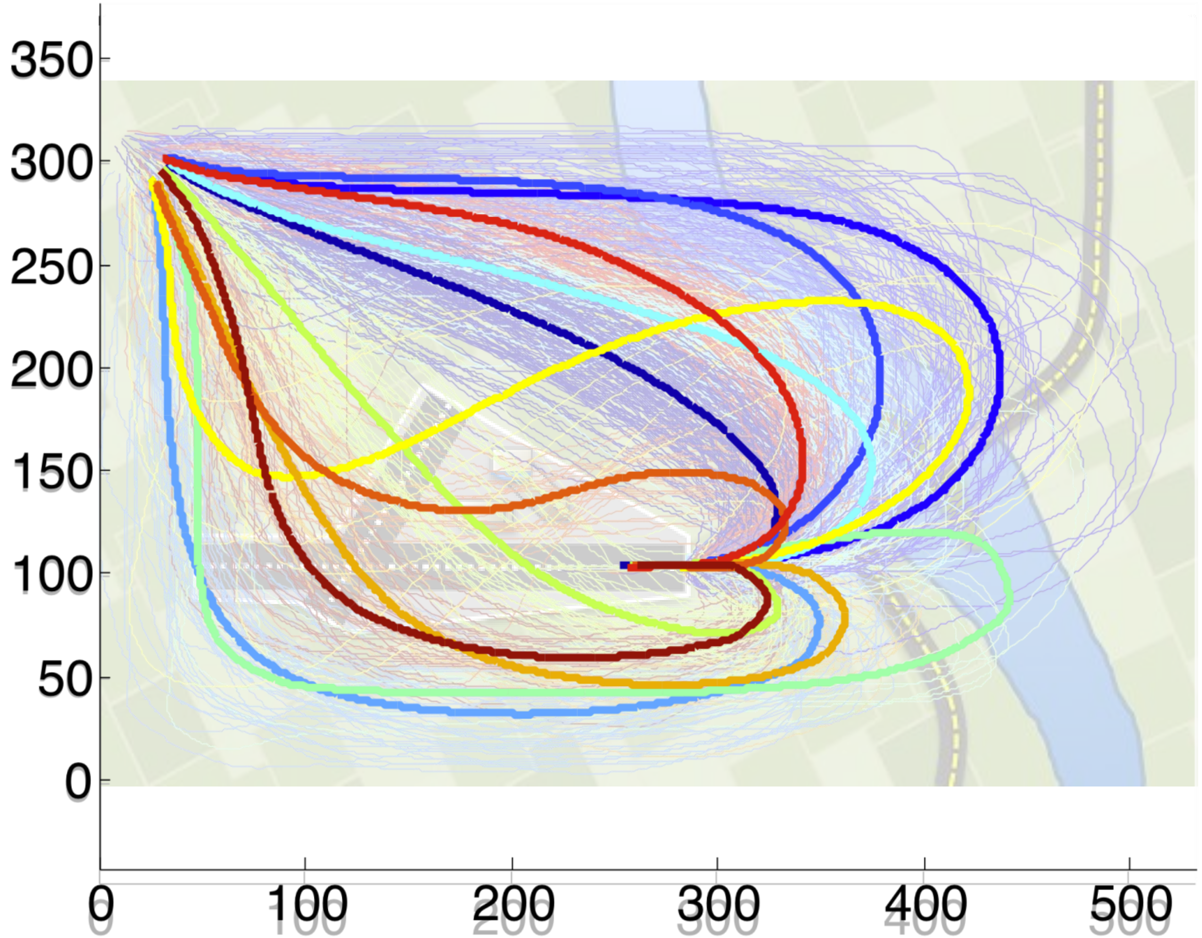
Sort by? High Score Top: 10 Select

Name	Planes Landed
wizbang_fll	9489
Tobbesuger	8853
7203644221	7761
Rhiannon	6215
lolblock	6210
tony	5963
spa	5651
AK	5073
anek	4808

[1]M. Stone et al., “Prospective memory in dynamic environments: Effects of load, delay, and phonological rehearsal.” *Memory*, 2001.

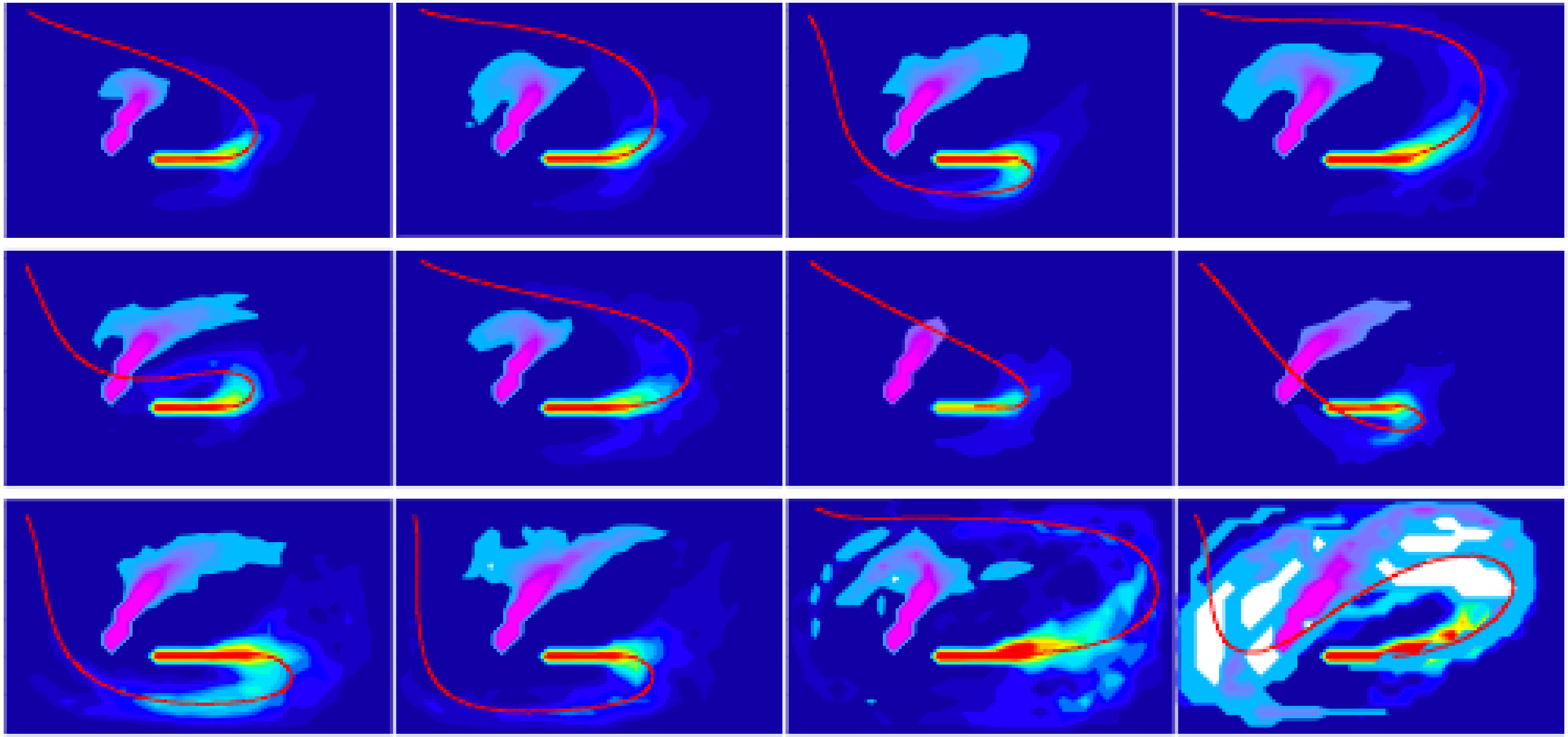
Modeling using gathered data

- Hypothesized hybrid model for controlled aircraft
- Data is supportive; clustering suggests discrete set of maneuvers used



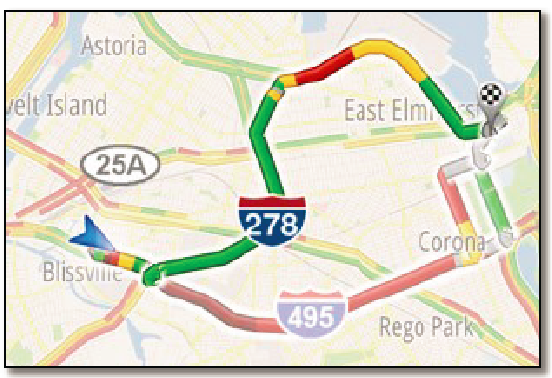
Modeling using gathered data

- Predict the maneuver given the airspace
  - Avoidance maneuvers
  - How people sequence moving objects

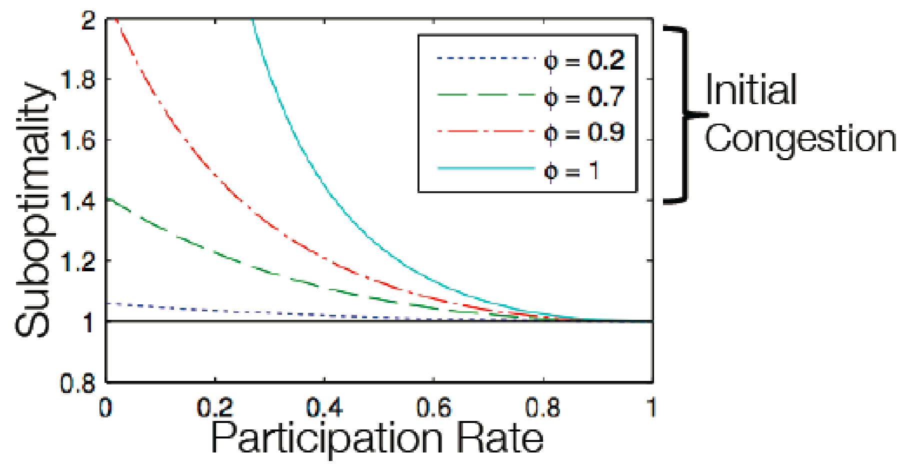


Analytics and Incentives

- **Design of visual big data analytics**
  - Encourage sustainable use of systems
  - Improve healthiness and well-being
- **Game theory to design incentives**
  - Connects groups with different objectives
  - Open problems in computational design



GPS-Traffic Display



[Aswani and Tomlin, ACC 2011; with Krichene, Zhou, Bayen 2013]

Conclusions and current work

- Reachability-inspired control of hybrid systems
  - Control law directly from the reachable set calculation
  - Automated controller synthesis for switched systems
    - Under sampling and quantization
  - Learning-based control inside reachable sets
    - Physics augmented with learning from data
- Current directions
  - Reachable set over-approximation methods
    - Control and “open loop” games allows for use of Fast Marching Methods
  - Capturing and using human behavior
  - Mechanisms in ATC (with Hamsa)
  - Stochastic hybrid systems: Models, Reachability methods, and Trajectory optimization
  - Mean-field game representations; principal-agent models (see poster by Balandat and Yang)

Learning and Control

Anil Aswani, Haomiao Huang, Zhengyuan Zhou, Claire Tomlin, Shankar Sastry

Department of Electrical Engineering and Computer Sciences
University of California at Berkeley

April 12, 2013

Results presented here have been supported by:
NSF CPS ActionWebs, ONR, AFOSR