Event-Based Optimization of Stochastic Systems

- A new optimization framework

Xi-Ren Cao

Shanghai Jiaotong University

Hong Kong University of Science and Technology

IEEE CSS Distinguished Lecture

August 15, 2013, Vancouver, Canada

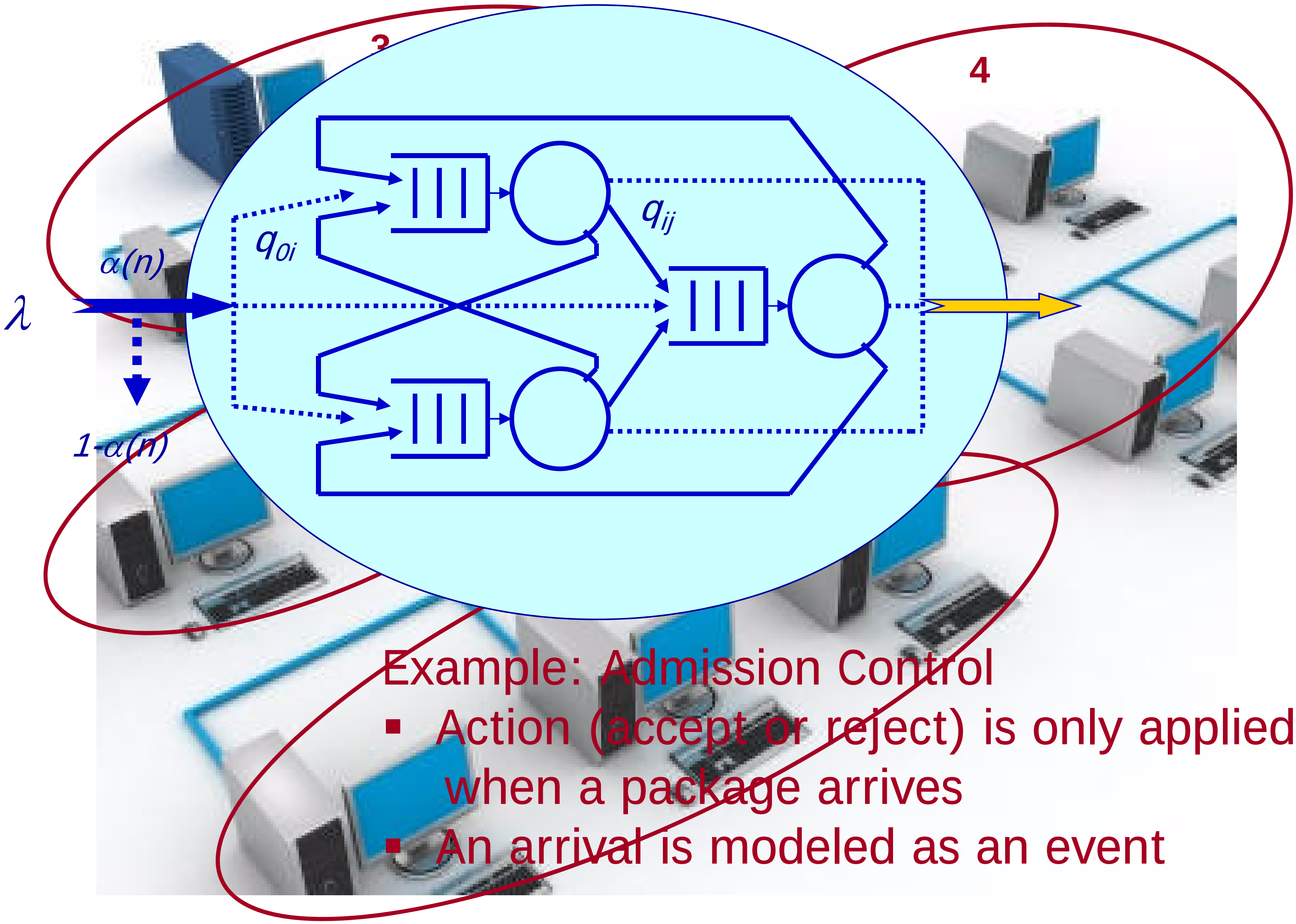
Contents

- Motivation
- Formulation of event-based optimization
- Intuitive solution
- Theoretical justification
  - Direct comparison based approach (vs dynamic programming)
- Applications and Examples
- Discussion

Motivation

- In many engineering, financial, and social problems, actions are only taken when some events happen
- State space is usually too large to handle

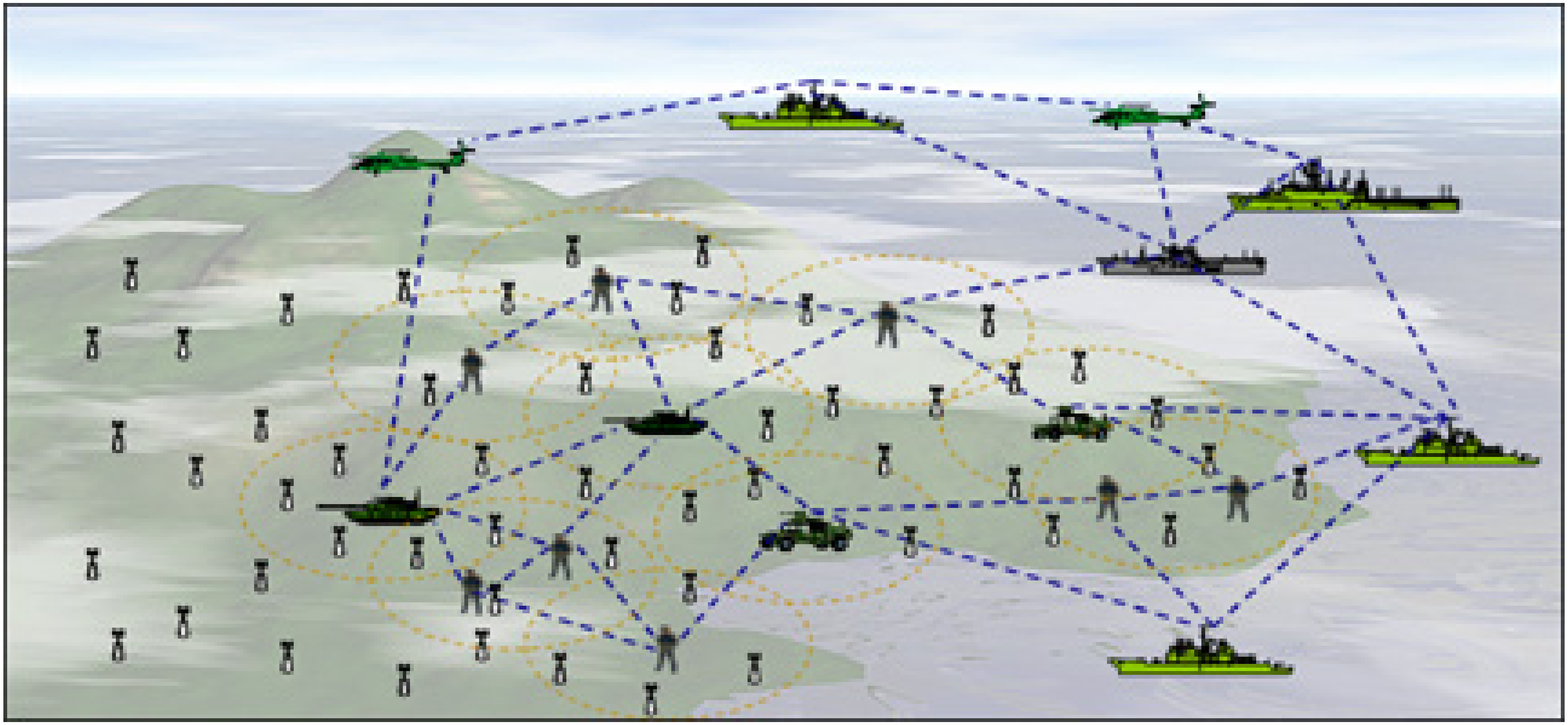
Network of Networks



Example: Admission Control

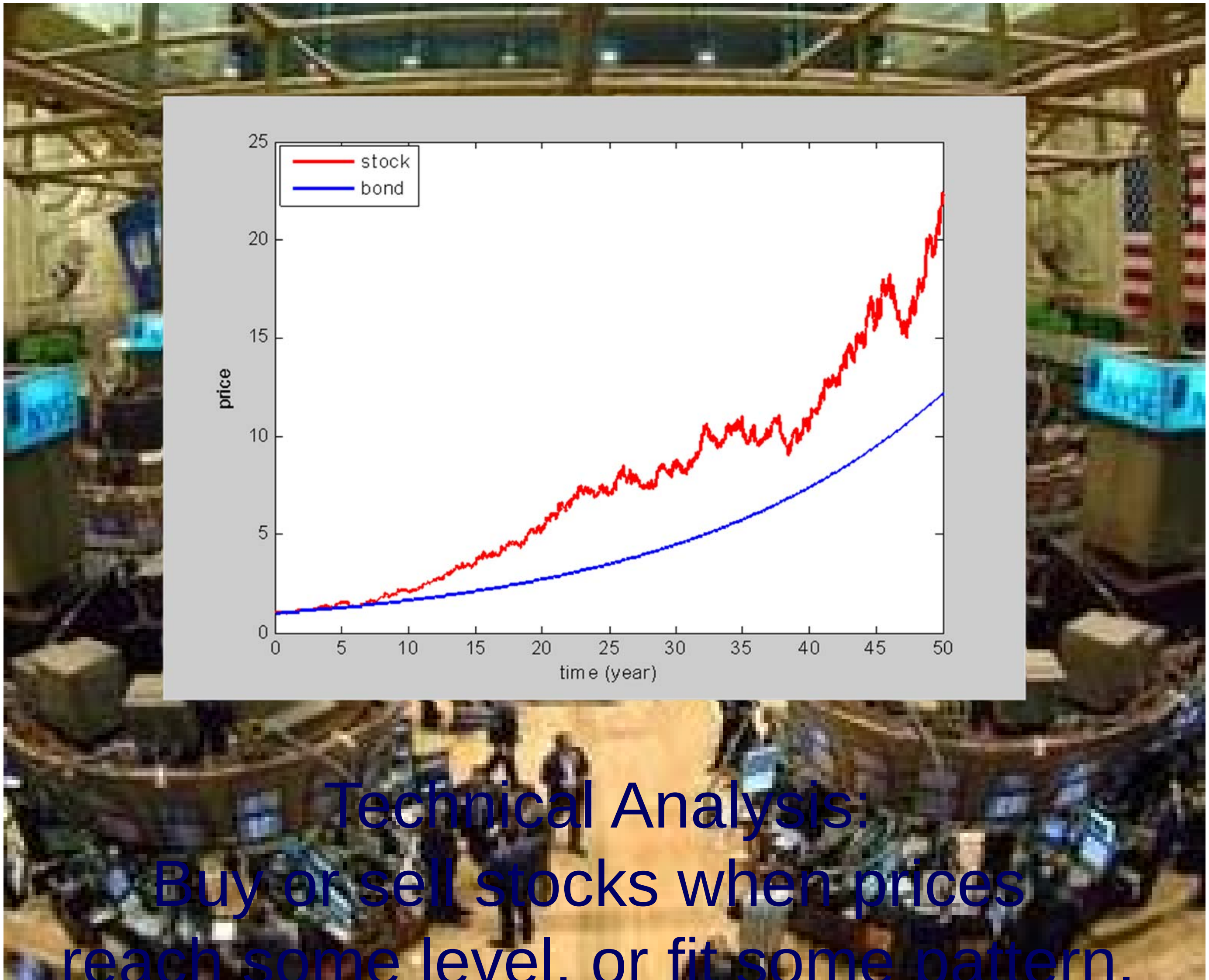
- Action (accept or reject) is only applied when a package arrives
- An arrival is modeled as an event

Sensor Networks



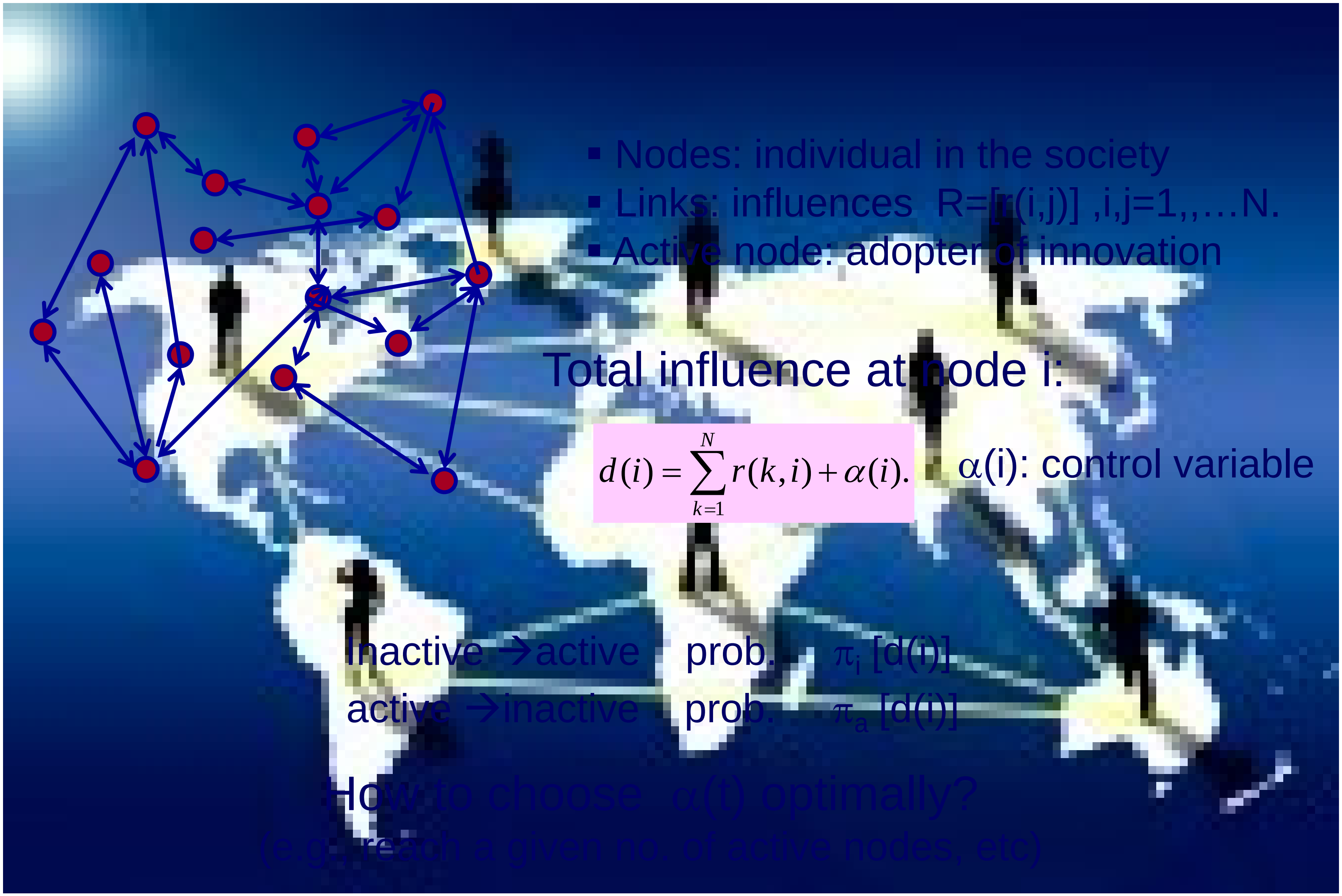
Actions are taken only when intruders are detected

Stock Market



Technical Analysis:
Buy or sell stocks when prices reach some level, or fit some pattern.

Innovation Spread on a Social Network



- Nodes: individual in the society
- Links: influences R=[r(i,j)] ,i,j=1,,…N.
- Active node: adopter of innovation

Total influence at node i:

$$d(i) = \sum_{k=1}^{N} r(k,i) + \alpha(i).$$

$\alpha(i)$: control variable

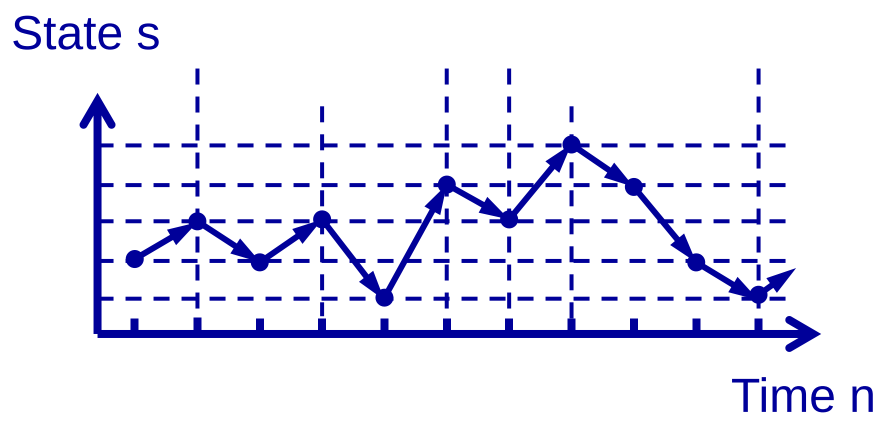
Inactive →active   prob. π_i [d(i)]

active →inactive   prob. π_a [d(i)]

How to choose $\alpha(t)$ optimally?

(e.g., reach a given no. of active nodes, etc)

Mathematical Formulation



- Markov chain $X=\{X_n, n=1,2,.....\}$, X_n in $S=\{1,2,...,M\}$.
- Transition prob. matrix $P=[p(j|i)]$, $i,j=1,..,M$
- Defined on probability space $\Omega=\{\omega:=\{X_n, n=1,...\}\}$

In probability theory, an event is a subset of Ω

Events

- An event: A set of transitions sharing common properties

$$e := \{\langle i, j \rangle : i, j \in S \text{ and } \langle i, j \rangle \text{ has common properties}\}$$

- Any ω in Ω can be viewed as a sequence of events;

  e.g., $\{X_n, X_{n+1}, X_{n+2}\} = \langle X_n, X_{n+1} \rangle, \langle X_{n+1}, X_{n+2} \rangle$

- Examples:
  - Admission: A customer arrives at a network

    $$e := \{\langle n, n_+ \rangle, \langle n, n \rangle\}$$
  - Sensor network:

    a tank crossed the border between two regions
  - Stock: price increases three days in a row
  - Social network:

    total influence of a node $\sum_{i=1}^{N} r(k,i)$ reaches some level
  - State aggregation, j in S, i in Ω

Formulation of EBO

- Event-based policy *d(e)=*α*.*
- Action α at event e, determines transition. prob.,

$$p^{\alpha}(j \mid i, e), \quad i, j \in S, \quad \langle i, j \rangle \in e.$$

- Events are observable, states may not

  We know $\langle i, j \rangle \in e.$ but may not know, *i* and *j*;

  we may control the prob. i→j in event e

- EBO: find a best policy

d = arg {max* η^{d}*: all d }*

$$\eta = \lim_{N \to \infty} \frac{1}{N} E\{\sum_{n=0}^{N-1} f(X_n) \mid X_0\} = \pi f = \sum_{i \in S} \pi(i) f(i)$$

Event sequence is NOT Markov!

Intuitive Solution

Markov Decision Process (MDP)

- At a state i, we apply an control action α, which controls reward and transition probabilities at i:

  $f(i,\alpha)$, $p^{\alpha}(j|i)$, α in A

- Choose action according to a policy d: $\alpha=d(i)$

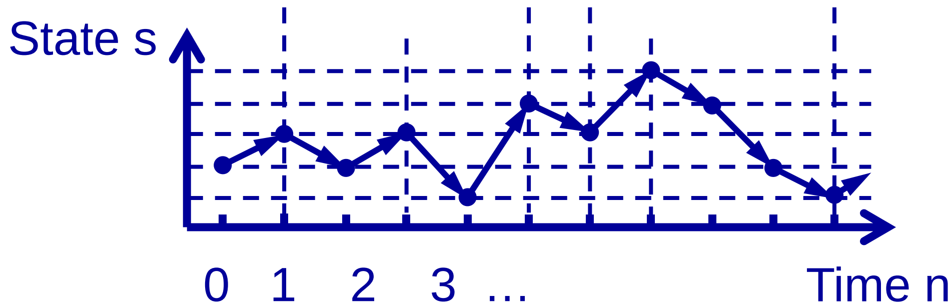
Every quantity depends on the policy applied to the system

Steady-state probability: $\pi^d=(\pi^d(1), \pi^d(2),..,\pi^d(M))$.

Transition Prob. matrix: $P^d=[p^{d(i)}(j/i)]_{i,j=1,..,M}$

System performance: $\eta^d=\pi^d f^d$

Intuition in Optimization



$$\eta_N^d = \frac{1}{N} E\{\sum_{n=0}^{N-1} f[X_n, d(X_n)] \mid X_0 = i\} = \frac{1}{N} Q_N^d(i,\alpha)$$

Q-factor: $$Q_N^d(i,\alpha) = E\left[f(X_0, A_0) + \sum_{l=1}^{N} f(X_l, d(X_l)) \middle| X_0 = i, A_0 = \alpha \right]$$

How about to choose

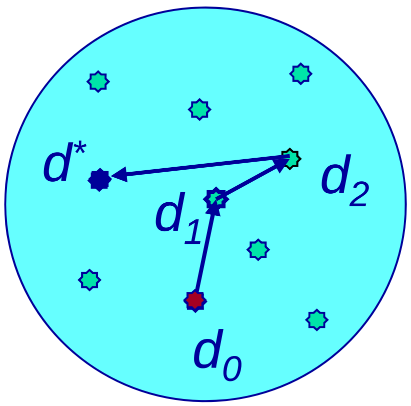
$$h(i) := \arg\max_{\alpha \in A} Q_N^d(i,\alpha)$$?

Proposition1: Suppose N is large enough, then

i) $\eta^h \geq \eta^d$

ii) if h=d, then d is optimal

i) Policy Iteration



Repeating the improvement Procedure, we get to an optimal Policy iteratively

$\eta^{d0} < \eta^{d1} < \eta^{d2} < \ldots < \eta^{d*}$

ii) Optimality (HJB) equation:

$$Q_N^d(i, d(i)) = \max_{\alpha \in A} Q_N^d(i, \alpha)$$

Does the same intuition work for EBO?

$$Q_N^d(e,\alpha) = E\left[f(X_0, A_0) + \sum_{l=1}^{N} f(X_l, d(X_l)) \middle| E_0 = e, A_0 = \alpha \right]$$

$$h(e) := \arg\max_{\alpha \in A} Q_N^d(e, \alpha)$$

Under what conditions, we have

i) $\eta^h \geq \eta^d$

ii) if h=d, then d is optimal

?

Which lead to

policy iteration and optimality equation for EPO

Analysis

MDP: Dynamic programming

EBO: Not Markov, DP does not work

⟹ Direct comparison

Technical Preparation:

Results hold after adding a constant to Q, redefine

$$Q_N^d(e,\alpha) = E\left[f(X_0, A_0) + \sum_{l=1}^{N} \{ f(X_l, d(X_l)) - \eta^d \} \middle| E_0 = e, A_0 = \alpha \right]$$

Set

$$Q^d(e,\alpha) = \lim_{N\to\infty} Q_N^d(e,\alpha)$$

Similarly, redefine

$$Q_N^d(i,\alpha) = E\left[f(X_0, A_0) + \sum_{l=1}^{N} \{ f(X_l, d(X_l)) - \eta^d \} \middle| X_0 = i, A_0 = \alpha \right]$$

Set $Q^d(i,\alpha) = \lim_{N\to\infty} Q_N^d(i,\alpha)$

If we choose $h(i) := \arg\max_{\alpha \in A} Q^d(i,\alpha)$ then $\eta^h \geq \eta^d$

Optimality equation $Q^d(i, d(i)) = \max_{\alpha \in A} Q^d(i,\alpha)$

$$g^d(i) = \lim_{N\to\infty} E\{\sum_{n=1}^{N} [f(X_n, d(X_n)) - \eta^d] \mid X_1 = i\}. \qquad g^d = (g^d(1), \ldots, g^d(M))^T$$

Then $$Q^d(i,\alpha) = f(i,\alpha) + \sum_j p^\alpha(j \mid i) g^d(j)$$

⟹ *Poisson equation:* $(I - P^d)g^d + \eta^d e = f^d.$

Results take the form:

$$h(i) := \arg\max_{\alpha \in A} \{ f(i,\alpha) + \sum_j p^\alpha(j \mid i) g^d(j) \} \qquad \eta^h \geq \eta^d$$

Optimality equation:

$$f(i, d(i)) + \sum_j p^{d(i)}(j \mid i) g^d(j) = \max_{\alpha \in A} \{ f(i,\alpha) + \sum_j p^\alpha(j \mid i) g^d(j) \}$$

MDP with Direct Comparison

- Theoretical Justification

Poisson equation: $(I - P^d)g^d + \eta^d e = f^d.$

Performance difference formula: *any two policies h and d:*

$$\eta^h - \eta^d = \pi^h\{(P^h g^d + f^h) - (P^d g^d + f^d)\}$$

$\pi^h > 0$ ➔ *if* $P^h g^d + f^h \geq P^d g^d + f^d$ *then* $\eta^h \geq \eta^d$

Policy iteration and HJB optimality equation:

Let P*, η* be an optimal policy, then

$$P^* g^* + f^* = \max_{h \in \mathbf{D}}\{P^h g^* + f^h\}$$

EBO theory with Direct Comparison (I)

Transition probability under event-based policy d :

$$p^{d}(j\mid i)=\sum_{e}\pi^{d}(e\mid i)\,p^{d(e)}(j\mid i,e)$$

Direct Comparison:

performance difference formula under policy h and d:

$$\eta^{h}-\eta^{d}=\pi^{h}[(P^{h}g^{d}+f^{h})-(P^{d}g^{d}+f^{d})]$$

$$=\sum_{e}\pi^{h}(e)\left\{\sum_{i}\pi^{h}(i\mid e)\left([\sum_{j}[p^{h(e)}(j\mid i,e)-p^{d(e)}(j\mid i,e)]g^{d}(j)+[f(j,h(e))-f(j,d(e))]\right)\right\}$$

$$=\sum_{e}\pi^{h}(e)\left\{Q^{h,d}(e,h(e))-Q^{h,d}(e,d(e))\right\}$$

where

$$Q^{h,d}(e,\alpha)=\sum_{i}\pi^{h}(i\mid e)[f(i,\alpha)+\sum_{j}p^{\alpha}(j\mid i,e)g^{d}(j)]$$

if $Q^{h,d}(e,h(e))\geq Q^{h,d}(e,d(e))\quad\forall e,$ then $\eta^{h}\geq\eta^{d}$

EBO theory with Direct Comparison (II)

if $Q^{h,d}(e,h(e)) \geq Q^{h,d}(e,d(e))\ \forall e,$ then $\eta^h \geq \eta^d$ → Almost Useless!

Need to solve for $\pi^h(i|e)$!

Condition for EBO Optimality:

$$\pi^h(i|e) = \pi^d(i|e), \quad \forall i, \forall e, \forall h, d.$$

⟹

$$Q^{h,d}(e,\alpha) = Q^d(e,\alpha) = \sum_i \pi^d(i,e)[f(i,\alpha) + \sum_j p^\alpha(j|i,e) g^d(j)]$$

⟹ if $Q^d(e,h(e)) \geq Q^d(e,d(e)) \quad \forall e,$ then $\eta^h \geq \eta^d$

$Q^\alpha(e,\alpha)$ Can be estimated by $Q_N^\alpha(e,\alpha)$

EBO theory with Direct Comparison (III)

Under Condition: $\pi^{h}(i \mid e) = \pi^{d}(i \mid e), \quad \forall i, \forall e, \forall h, d.$

→ if $h(e) := \arg\max_{a \in A} Q^{d}(e, a), \ \forall e \in \mathrm{E} \quad \forall e,$ then $\eta^{h} \geq \eta^{d}$

Repeating the procedure: Policy iteration

→ EBO HJB Optimality Equation

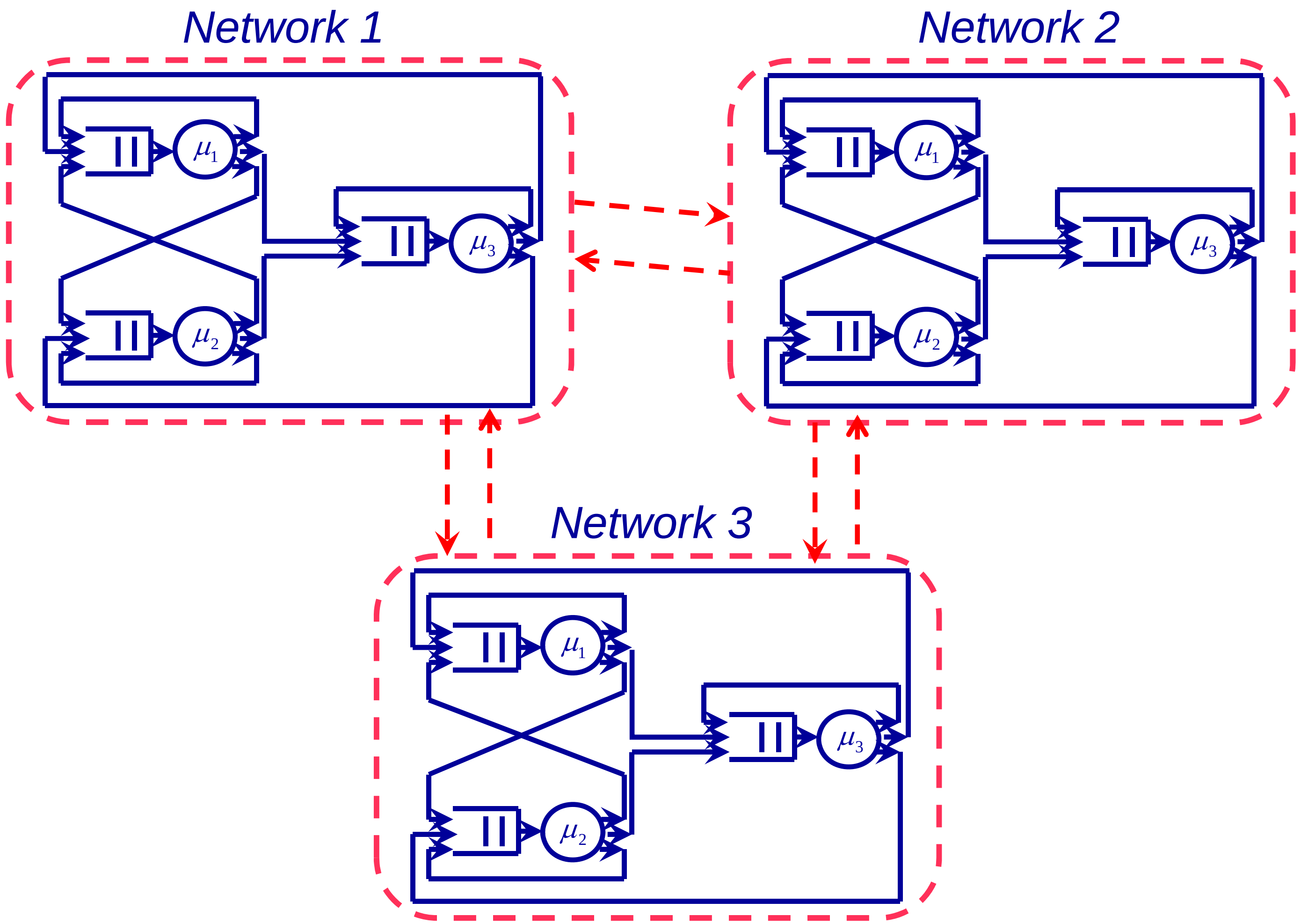
Let d* is an optimal event-based policy, then

$$Q^{d^*}(e, d^*(e)) = \arg\left(\max\left\{Q^{d^*}(e, \alpha), \forall \alpha\right\}\right).$$

Discussions

- Event formulation is wider than state aggregation
  - <i,j> contains "future information"
  - In continuous time, event cannot be modeled by expanding state space
- Not Markov, requires conditions, e.g., $\pi^h(i|e) = \pi^d(i|e)$
- If $\pi^h(i|I) = \pi^d(i|e)$, then e-optimal is global optimal
- When the condition $\pi^h(i|e) = \pi^d(i|e)$ does not hold, approximate methods can be developed, the error depends on $|\pi^h(i|e) - \pi^d(i|e)|$
- The direct comparison approach can be applied to some non-Markov processes and non-linear performance, where dynamic programming fails.
- More precise event formulation:
  Observable, controllable, and natural transition events

Example 1. Control networks of Networks



Control the probabilities of red traffic

$n_{i,j}$: Number of customers at jth server of network i
N_i: $= sum_j$ of $n_{i,j}$ Number of customers in network I
State: $(n_{1,1},\dots n_{1,k_1};\dots\dots n_{M,1},\dots n_{M,k_M})$
Aggregated state $(N_1,\dots.N_M)$

Event: a customer transits from one network to another and finds $(N_1,\dots.N_M)$

From Product-form solution:

$$\pi^h(i \mid e) = \pi^d(i \mid e), \qquad \forall i, \forall e, \forall h, d.$$

→ HJB equation and policy iteration exist

Example 2. in financial engineering

(Joint work with Dexin Wang)

- In stock market, the state is the stock prices, and a policy determines when and how much one has to buy and sell.
- Apparently, such decisions cannot change the price, thus the required condition $\pi^h(i|e) = \pi^d(i|e)$ holds automatically

Technical Analysis:

Actions are taken according to the market trends

eg., buy when loss x%

sell after rising three days

....

EBO provides a perfect tool in TA:

Modeling trends as events

Stock price equation:

$$dp_1(t) = \mu(t)p_1(t)dt + \sqrt{v(t)}p_1(t)dB^s(t).$$

where appreciation rate and volatility satisfy the following equations:

$$d\mu(t) = \kappa^\mu(\theta^\mu - \mu(t))dt + \xi^\mu dB^\mu(t).$$

$$dv(t) = \kappa^v(\theta^v - v(t))dt + \xi^v \sqrt{v(t)}dB^v(t).$$

Define a particular **EVENTS** $m = (m_1, m_2)$ as follows:

$$m_1(k) = Var\{\Delta y(k), \Delta y(k-1), \ldots, \Delta y(k-L+1)\},$$

and

$$m_2(k) = 100 * \frac{MACD(k) - EMA(MACD, 9)}{p_1(k)}$$

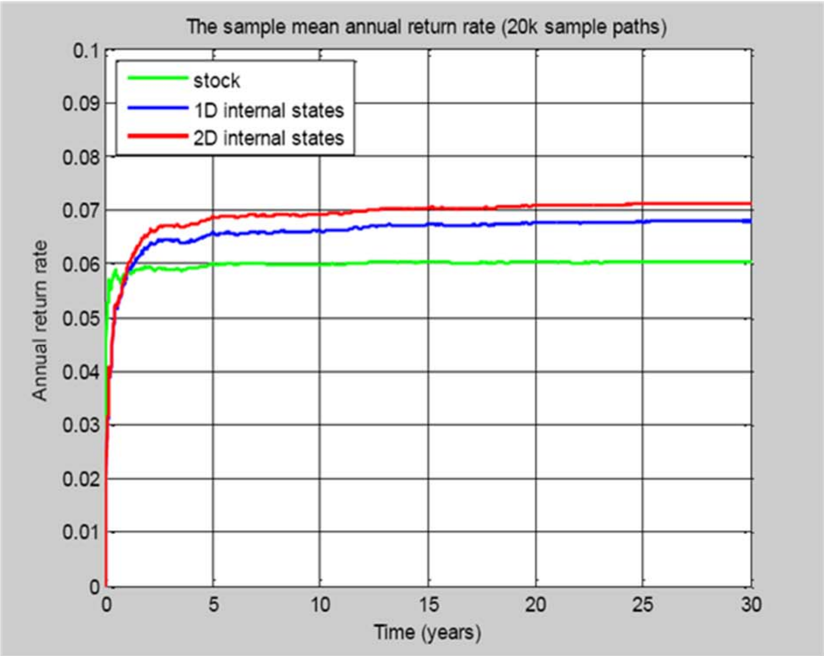
$$y(k) \;:=\; \ln(p_1(k)/p_1(0))$$

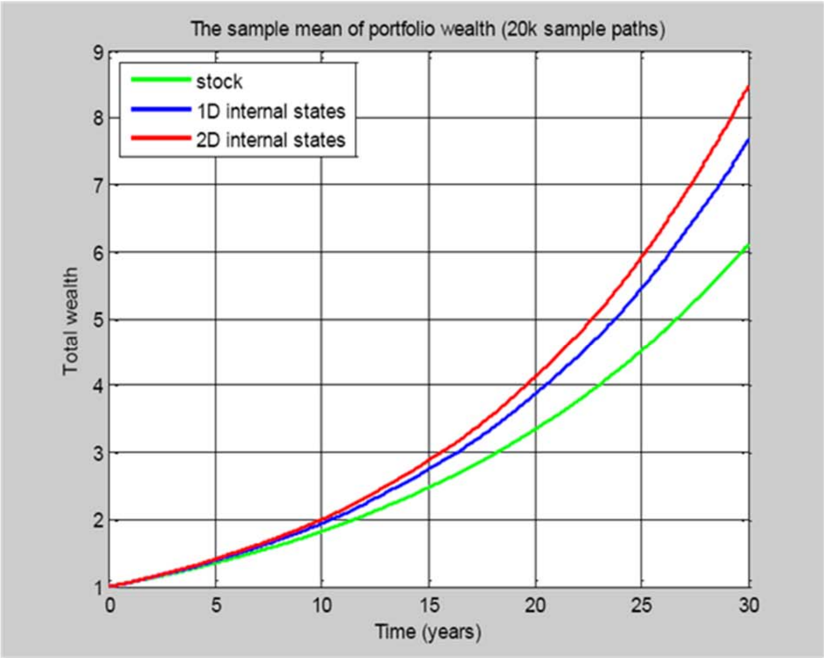
EMA: Exponential Moving Average $S_t = \alpha p_t + (1-\alpha)S_{t-1}$

MACD: Moving average convergence-divergence
=12 day EMA-26 day EMA

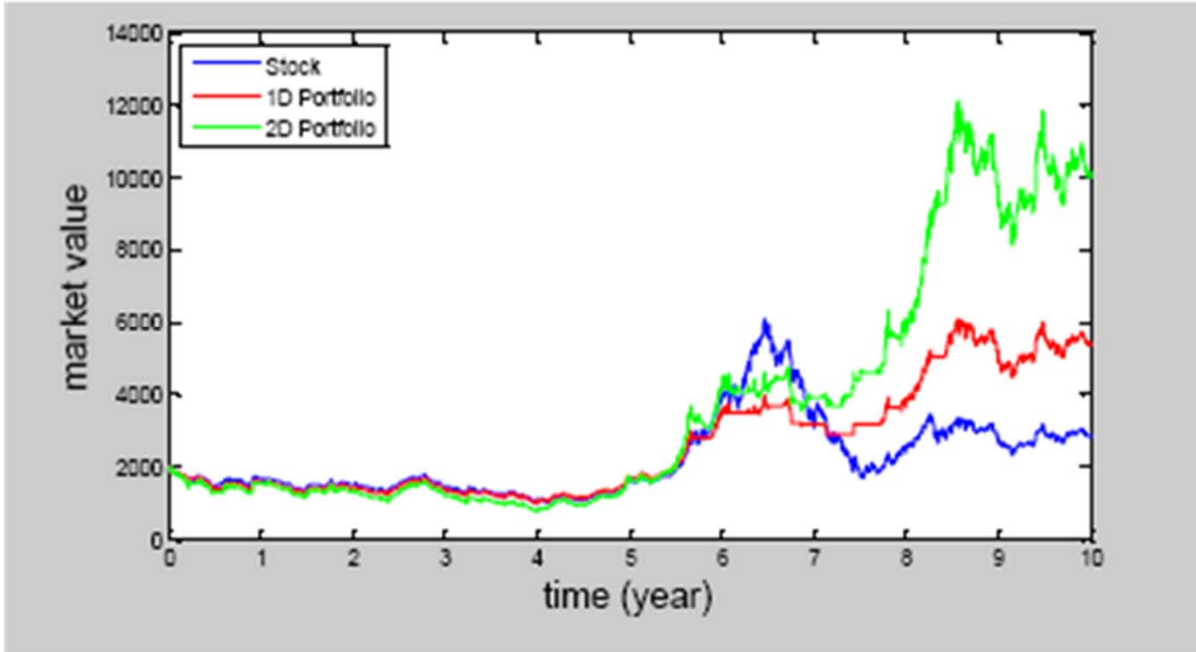
Results on a Stock

Strategy	Stock	Optimal 1D policy d^*	Optimal 2D policy d^{**}
Annual return	6.0%	6.8%	7.1%



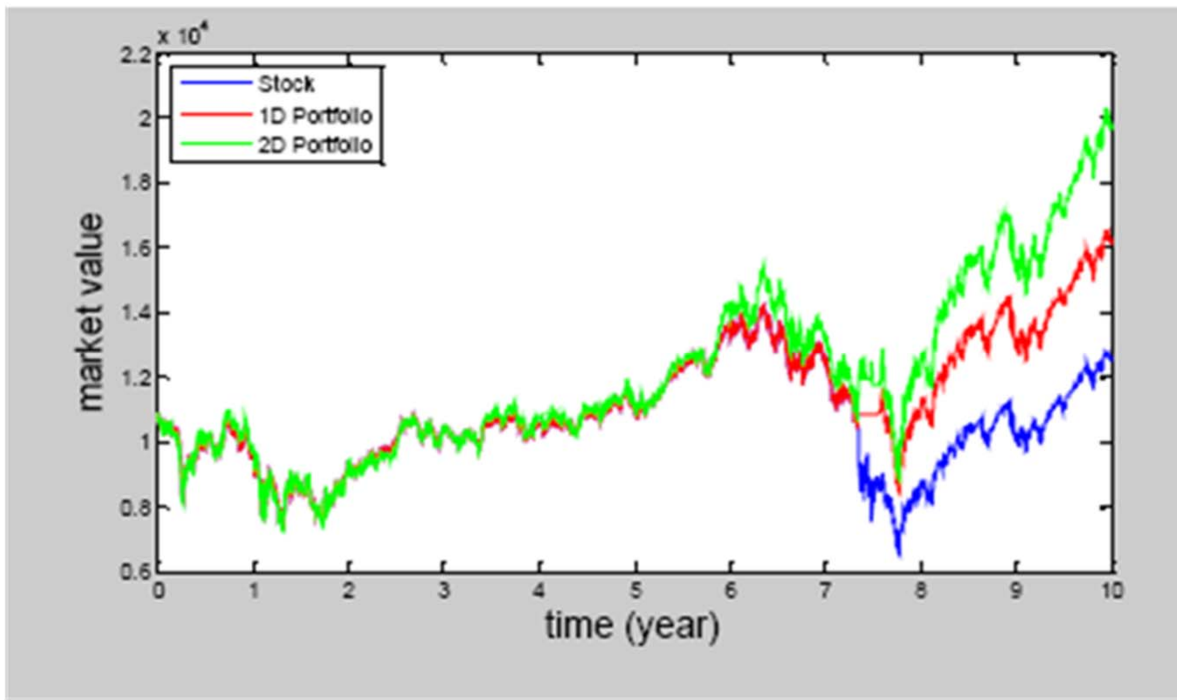


SSECI



14000
12000
10000
8000
6000
4000
2000
0
market value
Stock
1D Portfolio
2D Portfolio
0
1
2
3
4
5
6
7
8
9
10
time (year)

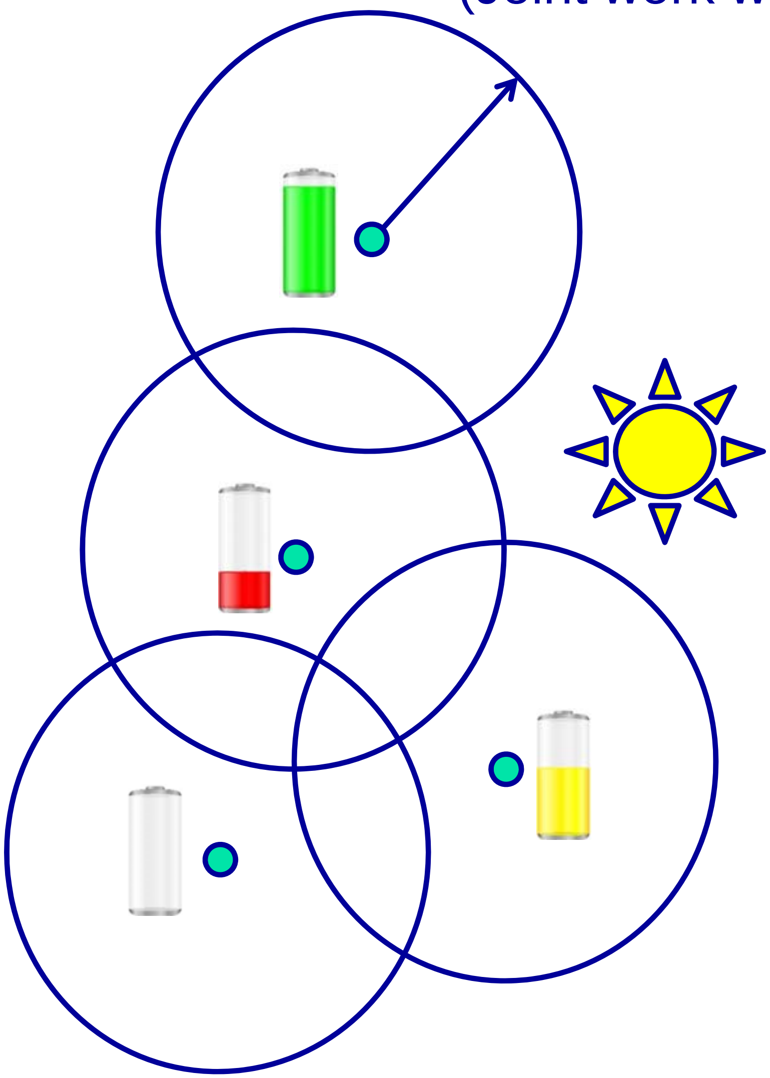
DJI



x 10^4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
market value
Stock
1D Portfolio
2D Portfolio
0
1
2
3
4
5
6
7
8
9
10
time (year)

Example 3. Activation policy optimization of wireless sensor networks (WSN)

(Joint work with Jia & Xia, Tsinghua Uni)



N sensors, Sensing radius, r_d
Detection probability, p
Discharging rate, R_d Solar charging rate, R_c

Events

1) A sensor just becomes fully charged and the ratio of active sensors is ((k-1)/n, k/n], k=1,…,n;
2) A sensor is about to be fully discharged and the ratio of active sensors is ((k-1)/n, k/n], k=1,…,n.

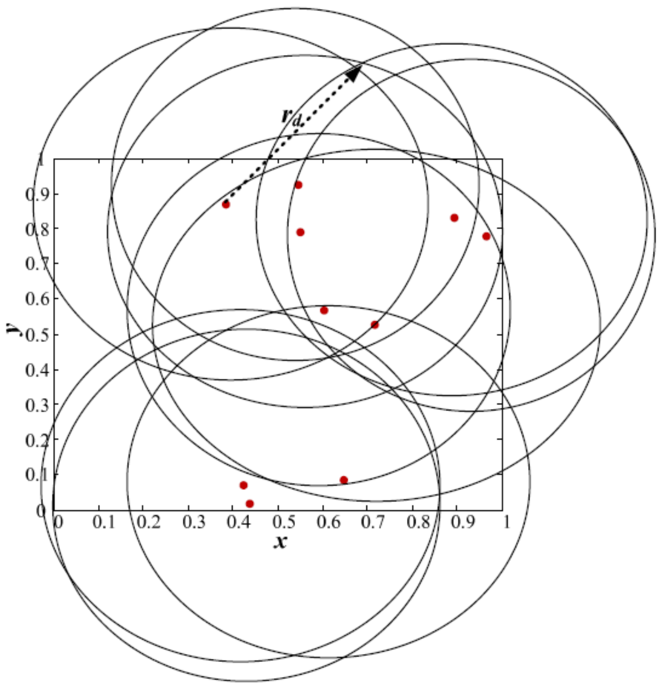
Actions

1) activate the current sensor.
2) activate a ready sensor

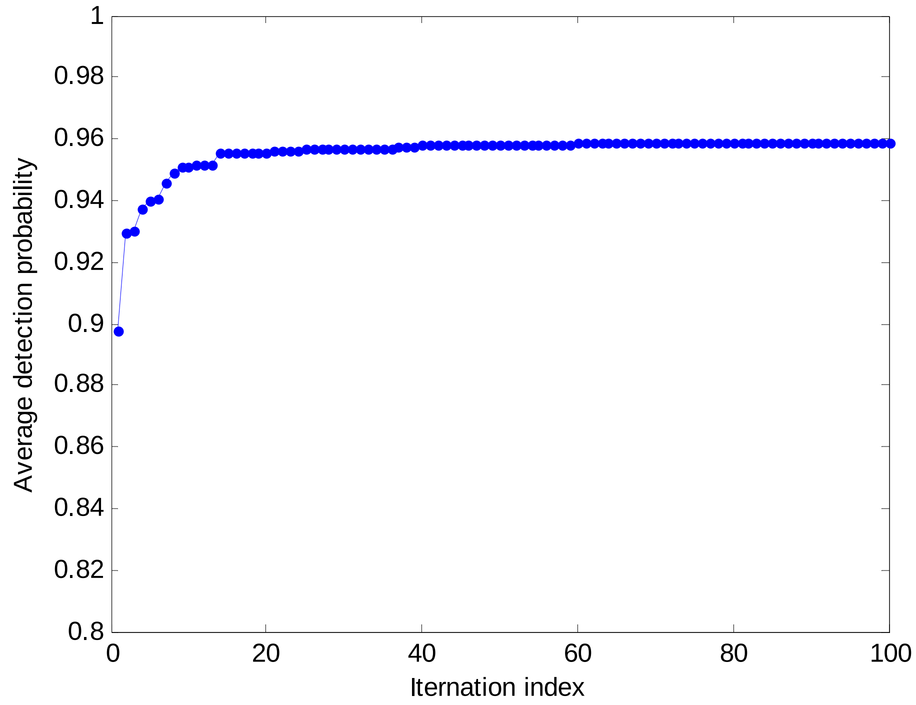
Objective: Maximize the average detection probability over the area in a given time.

Optimization results

- Parameter setting
  - $N=100$, $r_d=0.2$, $p=0.9$
  - $R_d=2$, $R_c=1$, $n=10$
  - Event space size: $2n=20$
  - While, state space size: 3^{100}

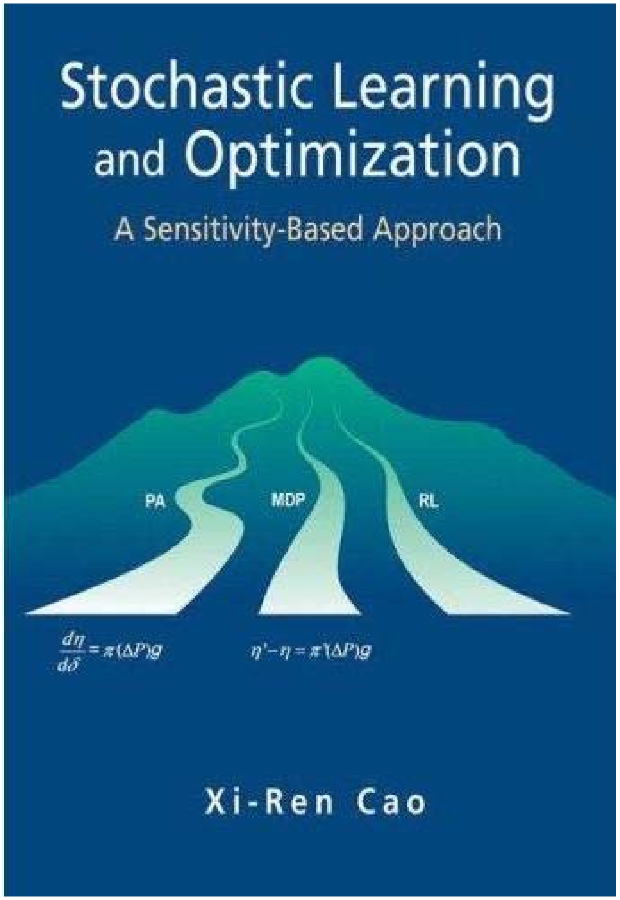


An example of sensor deployment over the target area



The curve of system performance w.r.t. number of iterations

Thank You!



Xi-Ren Cao:

Stochastic Learning and Optimization
- A Sensitivity Based Approach

9 Chapters, 566 pages, 119 Figures, 27 Tables,
212 homework problems

Springer   October 2007

X. R. Cao, L. Xia, and Q. S. Jia
Event-Based Optimization, A New Optimization Framework,
Discrete Event Dynamic Systems- Theory and Applications,
Special issue on Event Based Control, manuscript.