Optimal Control of Network Structure Growth

Dominik Thalmeier, Vicenç Gómez, Hilbert J. Kappen
SNN Adaptive Intelligence
Radboud University Nijmegen and Universitat Pompeu Fabra

Controlling network growth

Many real systems are networked systems.
- •Epidemic outbreaks
- •Knowledge transfer in science
- •Financial systems

Why controlling network growth?
•Those systems often do not work optimally
•It would be nice to have "social engineering methods" to make them work properly: hindering financial crisis or epidemic outbreaks
•The dynamics of networks, their resilience and also their controllability depend on the network topology
•The topology of a network is shaped during the growth of the network

Controlling network growth is difficult:
•Size of state space is growing super-exponentially
•Size of action space is often growing with the number of states

We propose to solve both problem by reducing the control problem to a sampling problem and learning an importance sampler which allows us to solve it approximately.

Stochastic optimal control

Formulation as MDP:

The state of the network e.g. Adjacency matrix	x
System transition given a controlling action *u*	$P(\mathrm{x}'\vert\mathrm{x},\mathrm{u})$
System transition given a control *u*	$\mathrm{r}(\mathrm{x},\mathrm{t})$

→ Minimize the total cost
$$\mathrm{r}(\mathrm{x},0)+\left\langle \sum_{t'=t+1}^{T}\mathrm{r}(\mathrm{x}_{t'},t')\right\rangle_{P(\mathrm{x}_{1:T}|\mathrm{x},\mathrm{u}(\cdot),\mathrm{t}=0)}$$

→ Bellman equation with value function *J*
$$J(\mathrm{x},\mathrm{t})=\min_{\mathrm{u}}\left(\mathrm{r}(\mathrm{x},\mathrm{t})+\langle J(\mathrm{x}',\mathrm{t}+1)\rangle_{P(\mathrm{x}'|\mathrm{x},\mathrm{u},\mathrm{t})}\right)$$
$$\mathrm{u}^*(\mathrm{x},\mathrm{t})=\mathrm{argmin}_{\mathrm{u}}\left(\mathrm{r}(\mathrm{x},\mathrm{t})+\langle J(\mathrm{x}',\mathrm{t}+1)\rangle_{P(\mathrm{x}'|\mathrm{x},\mathrm{u},\mathrm{t})}\right)$$

Problem:
The number of states (=possible networks) grows super-exponetially with the number of states!

Approximate control problem

We relax the problem:
→ Approximate MDP my a linear solvable MDP which can be formulated as a sampling problem

Solve for optimal transition-probabilities penalized by KL-divergence
$$\lambda\mathrm{KL}\left[\mathrm{u}(\mathrm{x}_{t+1:T}|\mathrm{x},\mathrm{t})\parallel \mathrm{p}(\mathrm{x}_{t+1:T}|\mathrm{x},\mathrm{t})\right]$$
$$+\mathrm{r}(\mathrm{x},\mathrm{t})+\left\langle \sum_{t'=t+1}^{T}\mathrm{r}(\mathrm{x}_{t'},t')\right\rangle_{\mathrm{u}(\mathrm{x}_{t+1:T}|\mathrm{x},\mathrm{t})}$$

New Bellman equation ca be solved by sampling
$$\phi(\mathrm{x}_{t+1:T}):=\exp\left(-\lambda^{-1}\sum_{t'=t+1}^{T}\mathrm{r}(\mathrm{x}_{t'},t')\right)$$
$$J^{\lambda}_{\mathrm{KL}}(\mathrm{x},\mathrm{t})=\mathrm{r}(\mathrm{x},\mathrm{t})-\lambda\log\langle\phi(\mathrm{x}_{t+1:T})\rangle_{\mathrm{p}(\mathrm{x}_{t+1:T}|\mathrm{x},\mathrm{t})}$$

To compute the action we maximize the approximate optimal value function
$$\mathrm{u}^*(\mathrm{x},\mathrm{t})\approx\mathrm{argmin}_{\mathrm{u}}\left(\mathrm{r}(\mathrm{x},\mathrm{t})+\langle J^{\lambda}_{\mathrm{KL}}(\mathrm{x}',\mathrm{t}+1)\rangle_{P(\mathrm{x}'|\mathrm{x},\mathrm{u},\mathrm{t})}\right)$$

Solving the sampling problem

Solve sampling problem by adaptive importance sampling using the **cross entropy method**.

Proposal distribution (with functional form of true optimal distribution):
$$\tilde{\mathrm{u}}_{\omega}(\mathrm{x}'|\mathrm{x},\mathrm{t})\propto\mathrm{p}(\mathrm{x}'|\mathrm{x})\exp\left(-\frac{\tilde{J}_{\mathrm{KL}}(\mathrm{x}',\omega(\mathrm{t}))}{\lambda}\right)$$

Where the value function is approximated with linear features of the network state.
$$\tilde{J}_{\mathrm{KL}}(\mathrm{x},\omega(\mathrm{t}))=\sum_{k}\omega_k(\mathrm{t})\psi^{\mathrm{t}}_k(\mathrm{x}).$$

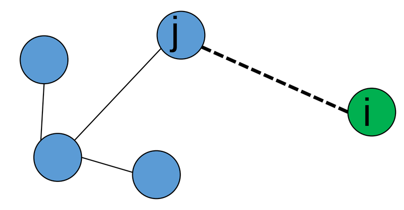
Computing the actual control

The value function $\tilde{J}_{\mathrm{KL}}$ learned by the Cross Entropy methods is a biased estimator of the true value function $J^{\lambda}_{\mathrm{KL}}(\mathrm{x},\mathrm{t})$.
We can obtain a consistent estimator by drawing samples:

$$\hat{\mathrm{u}}^*_{\mathrm{KL}}(\mathrm{x}'|\mathrm{x},\mathrm{t})=\left\langle\delta_{\mathrm{x}^{\mathrm{opt}}(\mathrm{t}+1),\mathrm{x}'}\right\rangle_{\mathrm{u}^*_{\mathrm{KL}}(\mathrm{x}_{t+1:T}|\mathrm{x},\mathrm{t})}$$
$$J^{\lambda}_{\mathrm{KL}}(\mathrm{x}',\mathrm{t}+1)\sim-\log\left(\frac{\hat{\mathrm{u}}^*_{\mathrm{KL}}(\mathrm{x}'|\mathrm{x},\mathrm{t})}{\mathrm{p}(\mathrm{x}'|\mathrm{x},\mathrm{t})}\right)$$

Task: maximizing the h-index of a growing cascade

•**Uncontrolled dynamics**: at every time-step, a new node ***i*** arrives and forms connections to the already existing nodes



•Focus is on trees: a new node connects to one parent node ***j***.
•Generative model of trees in which the probability of attach node *i* to parent node *j* is a combination three features: age τ, degree α and root bias β.
•Model is optimized using conversation trees from the site Slashdot.

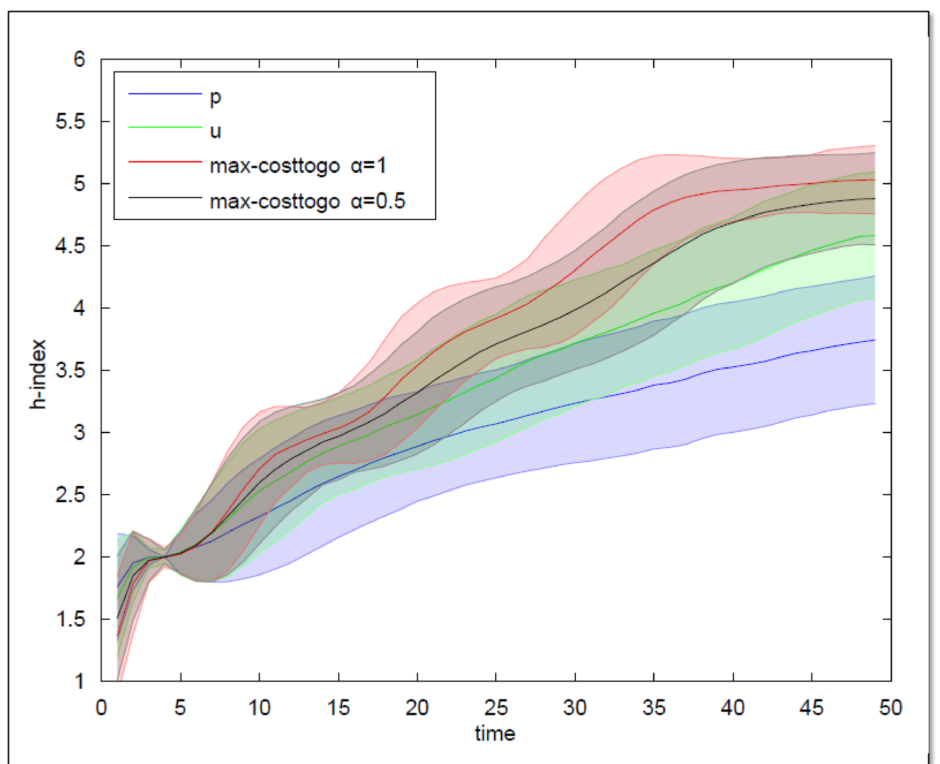
$$p(x'_i=j|x;\theta=(\alpha,\tau,\beta))=\frac{1}{Z_i}\left(k_j\alpha+\delta_{j,0}\beta+\tau^{i-j}\right)$$
$$-\log\mathcal{L}(\Pi;\theta)=-\sum_{k=1}^{N}\sum_{t=2}^{|x_{(k)}|}\log p({x_{(k)}}_t|[x_{(k)}];\theta)$$

•Interaction with the user is modelled by highlighting a node.
•**Simple model of user behavior**: a user picks the highlighted node *u* with probability $p'=\alpha/(1+\alpha)$, otherwise the user follows the dynamical model *p*. This gives the interaction model:

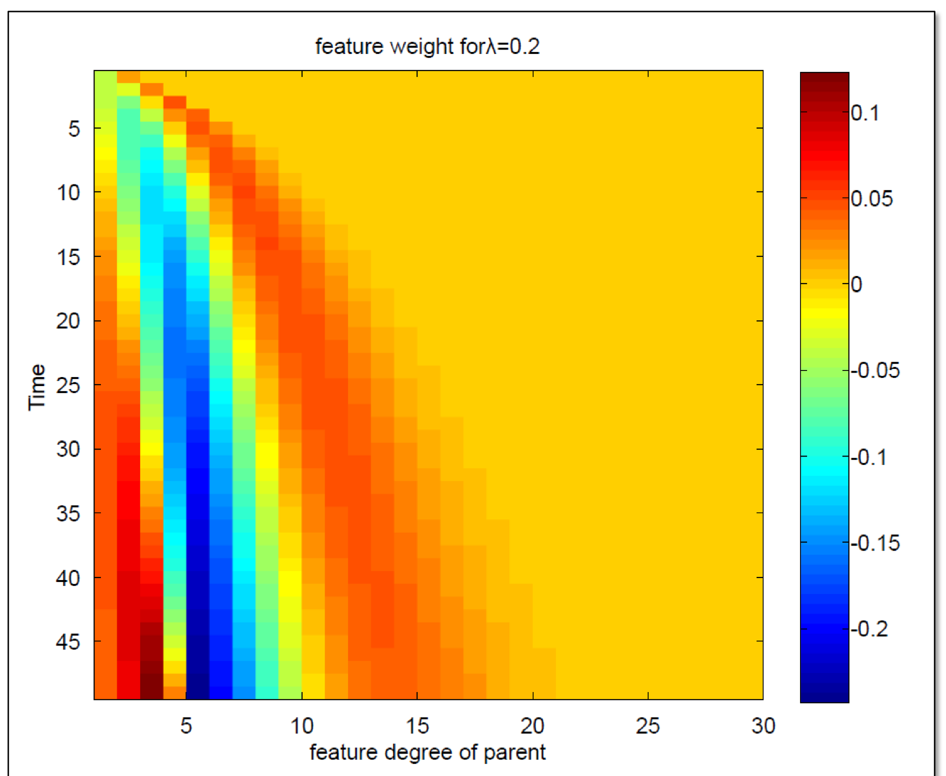
$$P(\mathrm{x}'|\mathrm{x},\mathrm{u})$$

•**State cost function r(x):** the **h-index** of a discussion tree, a topological measure previously linked to the controversy of a discussion
→ This is a non-trivial control problem: it is not possible to greedily optimize the action, as for most situations there is no action which immediately changes the h-index

Results:



We show the h-index over time, for different scenarios:
•Blue: uncontrolled dynamics – no node is highlighted
•Green: optimal solution for KL-regularized control directly on the transition probabilities
•**Black** and red: Dynamics with node highlighting



Learned Feature weights.
As features we used the degree of the parent nodes after the child attaches.

Blue means that new nodes are attracted to parents with this degree (low effective cost), while red is for repulsion (high effective cost).

References

E. Todorov. Efficient computation of optimal actions. *PNAS*, 106(28):11478–11483, 2009.

H. J. Kappen. Linear theory for control of nonlinear stochastic systems. *Physical Review Letters*, 95(20):200201, 2005.

V. Gómez, H. J. Kappen, N. Litvak, and A. Kaltenbrunner. A likelihood-based framework for the analysis of discussion threads. *World Wide Web*, 16(5-6):645–675, 2013.

V. Gómez, A. Kaltenbrunner, and V. López. Statistical analysis of the social network and discussion threads in Slashdot. In *Proceeding of the 17th international conference on World Wide Web*, pages 645–654. ACM, 2008.

Acknowledgements

This project is supported by the European Commission through FP7 Marie Curie ITN project "NETT" (Grant Nr. 28914)

Full paper accepted at Journal of Physics A
Preprint at: arXiv:1606.07777