# Mean Field Methods for Computer and Communication Systems

Jean-Yves Le Boudec EPFL Network Science Workshop Hong Kong 25-27 July 2012

#### Contents

- 1. Mean Field Interaction Model
- 2. Convergence to ODE
- 3. Finite Horizon: Fast Simulation and Decoupling assumption
- 4. Infinite Horizon: Fixed Point Method and Decoupling assumption

# MEAN FIELD INTERACTION MODEL

# **Mean Field**

A *model* introduced in Physics

interaction between *particles* is via distribution of states of all particle

#### An *approximation* method for a large collection of particles

► assumes *independence* in the master equation

#### Why do we care in information and communication systems ?

- Model interaction of many objects:
- Distributed systems, communication protocols, game theory, selforganized systems

### **A Few Examples Where Applied**

- [1] L. Afanassieva, S. Popov, and G. Fayolle. Models for transporation networks. *Journal of Mathematical Sciences*, 1997 – Springer.
- [2] F. Baccelli, A. Chaintreau, D. De Vleeschauwer, and D. R. McDonald. Http turbulence, May 2004.
- [3] F. Baccelli, M. Lelarge, and D. McDonald. Metastable regimes for multiplexed tcp flows. In *Proceedings of the*
- [5] M.-D. Bordenave, Charles and A. Proutiere. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. arXiv:math/0701363v2.
- [11] S. Kumar and L. Massoulié. Integrating streaming and filetransfer internet traffic: Fluid and diffusion approximations. MSR-TR-2005-160.
- [16] Y. M. Suhov and N. D. Vvedenskaya. Dobrushin's meanfield approximation for a queue with dynamic routing. *Markov Processes and Related Fields*, 3(4):493–526, 1997.
- [17] P. Tinnakornsrisuphap and A. M. Makowski. Limit behavior of ecn/red gateways under a large number of tcp flows. In Proceedings IEEE INFOCOM 2003, The 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, San Franciso, CA, USA, March 30 - April 3 2003.



5

# **Mean Field Interaction Model**

- Time is discrete (this talk) or continuous
- N objects, N large
   Object n has state X<sub>n</sub>(t)
   (X<sup>N</sup><sub>1</sub>(t), ..., X<sup>N</sup><sub>N</sub>(t)) is Markov
- Objects are observable only through their state

"Occupancy measure"
 M<sup>N</sup>(t) = distribution of
 object states at time t

# **Example: 2-Step Malware**

- Mobile nodes are either
  - ▶ `S' Susceptible
  - `D' Dormant
  - ► `A' Active
- Time is discrete
- Transitions affect 1 or 2 nodes
- State space is finite
  = {`S', `A',`D'}
- Occupancy measure is M(t) = (S(t), D(t), A(t)) with S(t) + D(t) + A(t) = 1
- S(t) = proportion of nodes in state `S'

```
[Benaïm and Le Boudec(2008)]
```

- 1. Recovery
  - ► D -> S
- 2. Mutual upgrade
  - ▶ D + D -> A + A
- 3. Infection by active
  - ▶ D + A -> A + A
- 4. Recovery
  - ► A -> S
- 5. Recruitment by Dormant
  - ▶ S + D -> D + D

Direct infection

- ► S -> D
- 6. Direct infection
  - ► S -> A

# **2-Step Malware – Full Specification**



#### Simulation Runs, N=1000 nodes



 $\beta = 0.01, \delta_A = 0.005, \delta_D = 0.0001, \alpha_0 = \alpha = 0.0001, h = 0.3, r = 0.1, \lambda = 0.0001$ 

#### Sample Runs with N = 1000



 $\beta = 0.01, \delta_A = 0.005, \delta_D = 0.0001, \alpha_0 = \alpha = 0.0001, h = 0.3, r = 0.1, \lambda = 0.0001$ 

# **The Importance of Being Spatial**



- Mobile node state = (c, t)
   c = 1 ... 16 (position)
   t ∈ R<sup>+</sup> (age of gossip)
- Time is continuous
- Occupancy measure is  $F_c(z,t)$  = proportion of nodes that at location *c* and have age  $\leq z$

[Age of Gossip, Chaintreau et al.(2009)]



# What can we do with a Mean Field Interaction Model ?

- Large *N* asymptotics, Finite Horizon
  - fluid limit of occupancy measure (ODE)
  - decoupling assumption (fast simulation)

#### Issues

- When valid
- How to formulate the fluid limit

- Large *t* asymptotic
  - Stationary approximation of occupancy measure
  - Decoupling assumption

#### Issues

► When valid



# 2. Convergence to ode

# To Obtain a Mean Field Limit we Must Make Assumptions about the Intensity *I(N)*

I(N) = (order of) expected number of transitions per object per time unit

A mean field limit occurs when we re-scale time by I(N) *i.e. one time slot*  $\approx I(N)$  i.e. we consider  $X^N(t/I(N))$ 

I(N) = O(1): mean field limit is in discrete time [Le Boudec et al (2007)]

I(N) = O(1/N): mean field limit is in continuous time [Benaïm and Le Boudec (2008)] (this talk)

## Intensity for this model is 1/N



#### Simulation Runs, N=1000 nodes

 $\beta = 0.01, \delta_A = 0.005, \delta_D = 0.0001, \alpha_0 = \alpha = 0.0001, h = 0.3, r = 0.1, \lambda = 0.0001$ 

In one time step, the number of objects affected by a transition is 0, 1 or 2; mean number of affected objects is O(1)There are *N* objects Expected number of transitions per time slot per object is  $O\left(\frac{1}{N}\right)$ 

9

### **The Mean Field Limit**

Under very general conditions (given later) the occupancy measure converges, in law, to a deterministic process, m(t), called the mean field limit

$$M^N\left(\frac{t}{I(N)}\right) \to m(t)$$

Finite State Space => ODE



# **Sufficient Conditions for Convergence**

- [Kurtz 1970], see also [Bordenav et al 2008], [Graham 2000]
- Sufficient condition verifiable by inspection:

▶ probabilities at every time slot have a limit when  $N \rightarrow \infty$ 

- [Benaïm and Le Boudec(2008), Ioannidis and Marbach(2009)]
  - Let W<sup>N</sup>(k) be the number of objects that do a transition in time slot k. Note that E (W<sup>N</sup>(k)) = NI(N), where I(N) <sup>def</sup>=intensity. Assume

 $\mathbb{E}\left(W^{N}(k)^{2}\right) \leq \beta(N) \quad \text{with} \quad \lim_{N \to \infty} I(N)\beta(N) = 0$ when I(N) = 1/N the condition is true as soon as Second moment of number of objects affected in one timeslot  $\leq$  a constant

Similar result when mean field limit is in discrete time [Le Boudec et al 2007]

### **Example: Convergence to Mean Field**



Number of transitions per time step is bounded by 2, therefore there is convergence to mean field The Mean field limite is an ODE One time step corresponds to  $\Delta t = 1/N$ 

# **Formulating the Mean Field Limit**



#### **Convergence to Mean Field**





For the finite state space case, most cases are verifiable by inspection of the model For the general state space, things may be more complex (fluid limit is not an ODE, e.g. [Chaintreau et al, 2009], [Gomez-Serrano et al, 2012]) E.La

# FINITE HORIZON : FAST SIMULATION AND DECOUPLING ASSUMPTION

3.

# **The Decoupling Assumption**

- Often used in analysis of complex systems
- Says that k objects are asymptotically mutually independent  $(k \text{ is fixed and } N \rightarrow \infty)$
- What is the relation to mean field convergence ?

# **The Decoupling Assumption**

Often used in analysis of complex systems

- Says that k objects are asymptotically mutually independent  $(k \text{ is fixed and } N \rightarrow \infty)$
- What is the relation to mean field convergence ?

[Sznitman 1991] [For a mean field interaction model: ]

Decoupling assumption

 $\Leftrightarrow$ 

 $\widetilde{M^N}(t)$  converges to a deterministic limit

Further, if decoupling assumption holds,  $m(t) \approx$  state proba for any arbitrary object

# The Two Interpretations of the Mean Field Limit

At any time t  $P(X_n(t) = A') \approx A\left(\frac{t}{N}\right)$   $P(X_m(t) = D', X_n(t) = A') \approx D\left(\frac{t}{N}\right) A\left(\frac{t}{N}\right)$ where (D, A, S) is solution of ODE

Thus for N = 1000 and simulation step k = 300:

- ▶ Prob (node *n* is dormant)  $\approx 0.48$
- ▶ Prob (node *n* is active)  $\approx 0.19$
- ▶ Prob (node *n* is susceptible)  $\approx 0.33$
- *m(t)* approximates both
- 1. the occupancy measure  $M^N(t)$
- 2. the state probability for one object at time *t*, drawn at random among *N*



## **Fast Simulation**

The evolution for one object as if the other objects had a state drawn randomly and independently from the distribution *m(t)* 

Is valid over finite horizon whenever mean field convergence occurs

Can be used to perform «fast simulation», i.e., simulate in detail only one or two objects, replace the rest by the mean field limit (ODE)

$$p_j^N(t|i) = P(X_n^N(t) = j | X_n^N(0)$$
$$= i)$$
$$p_j^N\left(\frac{t}{N}|i\right) \approx p_j(t|i)$$

where  $\vec{p}(t|i)$  is the (transient) probability of a continuous time nonhomogeneous Markov process d $\vec{r}(t|i) = \vec{r}(t|i)T A(\vec{rrr}(t))$ 

 $\frac{d}{dt}\vec{p}(t|i) = \vec{p}(t|i)^T A\big(\vec{m}(t)\big)$ 

Same ODE as mean field limit, with different initial condition  $\frac{d}{dt}\vec{m}(t) = \vec{m}(t)^T A(\vec{m}(t))$  $= F(\vec{m}(t))$ 

## We can fast-simulate one node, and even compute its PDF at any time



# INFINITE HORIZON: FIXED POINT METHOD AND DECOUPLING ASSUMPTION

4.

# **Decoupling Assumption in Stationary Regime**

Stationary regime = for large *t* 

Here:

- ▶ Prob (node *n* is dormant)  $\approx 0.3$
- ▶ Prob (node *n* is active) ≈ 0.6
- ▶ Prob (node *n* is susceptible) ≈ 0.1
- Decoupling assumption says distribution of prob for state of one object is  $\approx \vec{m}(t)$  with  $\frac{d\vec{m}(t)}{dt} = F(\vec{m}(t))$ 
  - We are interested in stationary regime, i.e we do  $F(\vec{m}) = 0$



# Example: 802.11 Analysis, Bianchi's Formula

ODE for mean field limit

$$\frac{dm_0}{d\tau} = -m_0 q_0 + \beta(\vec{m}) (1 - \gamma(\vec{m})) + q_K m_K \gamma(\vec{m})$$
$$\frac{dm_i}{d\tau} = -m_i q_i + m_{i-1} q_{i-1} \gamma(\vec{m}) \qquad i = 1, ..., K$$

$$\beta(\vec{m}) = \sum_{i=0}^{K} q_i m_i$$
  
$$\gamma(\vec{m}) = 1 - e^{-\beta(\vec{m})}$$

1



See [Benaim and Le Boudec , 2008] for this analysis

m<sub>i</sub> = proba one node is in

backoff stage I

802.11 single cell

 $\beta$ = attempt rate

 $\gamma$  = collision proba

$$m_{i} = \frac{1}{q_{i}} \frac{1}{\sum_{k=0}^{K} \frac{\gamma^{k}}{q_{k}}}$$
Bianchi's  
Fixed  
Point  
Equation  
[Bianchi 1998]  

$$\gamma = 1 - e^{-\beta}$$

$$\beta = \frac{\sum_{k=0}^{K} \gamma^{k}}{\sum_{k=0}^{K} \frac{\gamma^{k}}{q_{k}}}$$

 $\gamma^i$ 

## **Example Where Fixed Point Method Fails**



# When the Fixed Point Method Fails, Decoupling Assumption Does not Hold

In stationary regime,  $\vec{m}(t) = (D(t), A(t), S(t))$  follows the limit cycle

- Assume you are in stationary regime (simulation has run for a long time) and you observe that one node, say n = 1, is in state 'A'
- It is more likely that m(t) is in region R
  - Therefore, it is more likely that some other node, say n = 2, is also in state 'A'

Nodes are not independent – they are *synchronized* 



# **Example: 802.11 with Heterogeneous Nodes**



[Cho et al, 2010]

Two classes of nodes with heterogeneous parameters (restransmission probability)

Fixed point equation has a unique solution, but this is not the stationary proba

There is a limit cycle

# Where is the Catch ?

Decoupling assumption says that nodes m and n are asymptotically independent

There *is* mean field convergence for this example

But we saw that nodes may not be asymptotically independent

... is there a contradiction ?



The *decoupling assumption may not hold in stationary regime*, even for perfectly regular models

A correct statement is: conditionally independent given the value of the mean field limit m(t)

Positive Result 1 [e.g. Benaim et al 2008] : Decoupling Assumption Holds in Stationary Regime if mean field limit ODE has a unique fixed point to which all trajectories converge



#### Positive Result 2: In the Reversible Case, the Fixed Point Method Always Works

Definition Markov Process X(t) with transition rates q(i,j) is reversible iff

1. it is ergodic 2. p(i) q(i,j) = p(j) q(j,i) for some p

**Theorem 1.2 ([Le Boudec(2010)])** Assume some process  $Y^N(t)$  converges at any fixed t to some deterministic system y(t) at any finite time. Assume the processes  $Y^N$  are reversible under some probabilities  $\Pi^N$ . Let  $\Pi \in \mathcal{P}(E)$  be a limit point of the sequence  $\Pi^N$ .  $\Pi$  is concentrated on the set of stationary points S of the fluid limit y(t)

Stationary points = fixed points

If process with finite N is reversible, the stationary behaviour is determined only by fixed points.

# A Correct Method in Order to Make the Decoupling Assumptions

- 1. Write dynamical system equations in transient regime
- 2. Study the stationary regime of dynamical system
  - if converges to unique stationary point m\*
     then make fixed point assumption
  - else objects are coupled in stationary regime by mean field limit m(t)
- Hard to predict outcome of 2 (except for reversible case)

# Stationary Behaviour of Mean Field Limit is not predicted by Structure of Markov Chain

- $M^{N}(t)$  is a Markov chain on  $S^{N}=\{(a, b, c) \ge 0, a + b + c = 1, a, b, c multiples of 1/N\}$ 
  - *M<sup>N</sup>(t)* is ergodic and aperiodic, for any value of *h*

 $S^{N}$  (for N = 200) 0.9 0.8 0.7 0.6 Active 0.4 0.3 0.2 0.1 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 Dormant

Depending on *h*, there is or is not a limit cycle for *m(t)* 



## Conclusion

- Mean field models are frequent in large scale systems
- Validity of approach is often simple by inspection
  - Mean field is both
    - ODE for fluid limit
    - Fast simulation using decoupling assumption

Decoupling assumption
 holds at finite horizon; may
 not hold in stationary regime
 (except for reversible case)

Study the stationary regime of the ODE !

(instead of computing the stationary proba of the Markov chain)

#### Thank You ...

#### References

- [Baccelli et al.(2004)Baccelli, Lelarge, and McDonald] F. Baccelli, M. Lelarge, and D McDonald. Metastable regimes for multiplexed tcp flows. In Proceedings of the Forty-Second Annual Allerton Conference on Communication, Control, and Computing, Allerton House, Monticello, Illinois, USA, University of Illinois at Urbana-Champaign, October 2004.
- [Benaim and Le Boudec(2008)] M. Benaim and J.Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11-12):823–838, 2008.
- [Benaïm and Weibull(2003)] M. Benaïm and J. Weibull. Deterministic approximation of stochastic evolution. *Econometrica*, 71:873–904, 2003.
- [Benaim et al.(2006)Benaim, Hofbauer, and Sorin] M. Benaim, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions ii: Applications. 2006.
- [Bordenave et al.(2007)Bordenave, McDonald, and Proutiere] C. Bordenave, D. McDonald, and A. Proutiere. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. Arxiv preprint math/0701363, 2007.

[Bordenave et al.(2008)Bordenave, McDonald, and Proutiere] C. Bordenave, D. McDonald, and A. Proutiere. Performance of random medium access control, an asymptotic approach. In *Proceedings of the 2008 ACM SIGMETRICS international conference on Measurement and modeling* of computer systems, pages 1–12. ACM, 2008.

- [Buchegger and Le Boudec(2002)] S. Buchegger and J.-Y. Le Boudec. Performance analysis of the confidant protocol (cooperation of nodes - fairness in dynamic ad-hoc networks). In *Proceedings of MobiHoc'02*, June 2002.
- [Chaintreau et al.(2009)Chaintreau, Le Boudec, and Ristanovic] A. Chaintreau, J.Y. Le Boudec, and N. Ristanovic. The age of gossip: spatial mean field regime. In *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, pages 109–120. ACM, 2009.
- [Darling and Norris(2008)] RWR Darling and J.R. Norris. Differential equation approximations for Markov chains. *Probability surveys*, 5:37–79, 2008.
- [Deffuant et al.(2000)Deffuant, Neau, Amblard, and Weisbuch] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. Advances in Complex Systems, 3:87–98, 2000.
- [Ethier and Kurtz(2005)] Stewart N. Ethier and Thomas G. Kurtz. Markov Processes, Characterization and Convergence. Wiley, 2005.

[Gast et al.(2010)Gast, Gaujal, and Le Boudec] Nicolas Gast, Bruno Gaujal, and Jean-Yves Le Boudec. Mean field for Markov Decision Processes: from Discrete to Continuous Optimization. Technical Report arXiv:1004.2342v2, 2010.

[Gomez-Serrano et al, 2012] Gomez-Serrano J., Graham C. and Le Boudec J.-Y.

The Bounded Confidence Model Of Opinion Dynamics *Mathematical Models and Methods in Applied Sciences*, Vol. 22, Nr. 2, pp. 1150007-1--1150007-46, 2012.

- [Graham and Méléard(1994)] Carl Graham and Sylvie Méléard. Chaos hypothesis for a system interacting through shared resources. *Probab. Theory Related Fields*, 100(2):157–173, 1994. ISSN 0178-8051. doi: 10.1007/BF01199263. URL http://dx.doi.org/10.1007/BF01199263.
- [Graham and Méléard(1997)] Carl Graham and Sylvie Méléard. Stochastic particle approximations for generalized Boltzmann models and convergence estimates. Ann. Probab., 25(1):115–132, 1997. ISSN 0091-1798. doi: 10.1214/aop/1024404281. URL http://dx.doi.org/ 10.1214/aop/1024404281.
- [Ioannidis and Marbach(2009)] S. Ioannidis and P. Marbach. Absence of Evidence as Evidence of Absence: A Simple Mechanism for Scalable P2P Search. In *INFOCOM 2009, IEEE*, pages 576–584. IEEE, 2009.
- [Kelly(1991)] F.P. Kelly. Loss networks. The annals of applied probability, 1 (3):319–378, 1991.

- [Kurtz(1970)] T.G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1): 49–58, 1970.
- [Le Boudec(2010)] Jean-Yves Le Boudec. The Stationary Behaviour of Fluid Limits of Reversible Processes is Concentrated on Stationary Points. Technical Report arxiv:1009.5021.v2, 2010.
- [Le Boudec et al.(2007)Le Boudec, McDonald, and Mundinger] Jean-Yves Le Boudec, David McDonald, and Jochen Mundinger. A Generic Mean Field Convergence Result for Systems of Interacting Objects. In QEST'07, 2007.
- [McDonald(2007)] David McDonald. Lecture Notes on Mean Field Convergence, March 2007.
- [Sandholm(2006)] W.H. Sandholm. Population games and evolutionary dynamics. Unpublished manuscript, University of Wisconsin, 2006.
- [Sznitman(1991)] A.S. Sznitman. Topics in propagation of chaos. In P.L. Hennequin, editor, Springer Verlag Lecture Notes in Mathematics 1464, Ecole d'Eté de Probabilités de Saint-Flour XI (1989), pages 165Ű–251, 1991.

[Tembine et al.(2009)Tembine, Le Boudec, El-Azouzi, and Altman] Hamidou Tembine, Jean-Yves Le Boudec, Rachid El-Azouzi, and Eitan Altman. Mean Field Asymptotic of Markov Decision Evolutionary Games and Teams. In *Gamenets 2009*, 2009. Invited Paper.

[Tinnakornsrisuphap and Makowski(2003)] Peerapol Tinnakornsrisuphap and Armand M. Makowski. Limit behavior of ecn/red gateways under a large number of tcp flows. In Proceedings IEEE INFOCOM 2003, The 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, San Franciso, CA, USA, March 30 - April 3 2003.