Mean Field Methods for Computer and Communication Systems

Jean-Yves Le Boudec
EPFL
ACCESS Distinguished Lecture Series, Stockholm, May 28, 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
1

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, self-organized systems
Mean Field Interaction Model

- Time is discrete
- \( N \) objects, \( N \) large
- Object \( n \) has state \( X_n(t) \)
- \((X^1_N(t), ..., X^N_N(t))\) is Markov
- Objects are observable only through their state

"Occupancy measure"
\( M^N(t) = \) distribution of object states at time \( t \)
Mean Field Interaction Model

- Time is discrete
- $N$ objects, $N$ large
- Object $n$ has state $X_n(t)$
- $(X^N_1(t), ..., X^N_N(t))$ is Markov
- Objects are observable only through their state

- “Occupancy measure”
  $M^N(t) = \text{distribution of object states at time } t$

- **Theorem** [Gast (2011)]
  $M^N(t)$ is Markov

- Called “Mean Field Interaction Models” in the Performance Evaluation community
  [McDonald(2007), Benaïm and Le Boudec(2008)]
A Few Examples Where Applied


Example: 2-Step Malware

- Mobile nodes are either
  - `S` Susceptible
  - `D` Dormant
  - `A` Active
- Time is discrete
- Nodes meet pairwise (bluetooth)
- One interaction per time slot, \( I(N) = 1/N \); mean field limit is an ODE
- State space is finite
  \[ \{`S', `A',  `D'\} \]
- Occupancy measure is
  \[ M(t) = (S(t), D(t), A(t)) \] with
  \[ S(t)+ D(t) + A(t) =1 \]
  \[ S(t) = \text{proportion of nodes in state} \ `S' \]

[Benaïm and Le Boudec(2008)]

- Possible interactions:
  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`
Simulation Runs, $N=1000$ nodes

D(t)
Proportion of nodes
In state $i=1$

A(t)
Proportion of nodes
In state $i=2$

$\beta = 0.01, \delta_A = 0.005, \delta_D = 0.0001, \alpha_0 = \alpha = 0.0001, h = 0.3, r = 0.1, \lambda = 0.0001$
Sample Runs with \( N = 1000 \)

\[
\beta = 0.01, \delta_A = 0.005, \delta_D = 0.0001, \alpha_0 = \alpha = 0.0001, h = 0.3, r = 0.1, \lambda = 0.0001
\]
Example: TCP and ECN

- [Tinnakornsrisuphap and Makowski(2003)]

- Time is discrete, mean field limit is also in discrete time (iterated map)

- Similar examples:
  - HTTP Metastability [Baccelli et al.(2004)Baccelli, Lelarge, and McDonald]

At every time step, all connections update their state: I(N)=1
The Importance of Being Spatial

- Mobile node state = (c, t)
  c = 1 ... 16 (position)
  t ∈ R⁺ (age of gossip)

- Time is continuous, I(N) = 1

- Occupancy measure is
  \( F_c(z,t) = \text{proportion of nodes that at location c and have age } \leq z \)

[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
**Intensity $I(N)$**

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

- A mean field limit occurs when we re-scale time by $I(N)$
  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)]
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
Mean Field Limit
\( N = +\infty \)

Stochastic system
\( N = 1000 \)
Sufficient Conditions for Convergence

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

[Benaïm and Le Boudec(2008), Ioannidis and Marbach(2009)]

Let $W^N(k)$ be the number of objects that do a transition in time slot $k$. Note that $\mathbb{E}(W^N(k)) = NI(N)$, where $I(N) \equiv$ intensity. Assume

$$\mathbb{E}\left( W^N(k)^2 \right) \leq \beta(N) \quad \text{with} \quad \lim_{N \to \infty} I(N) \beta(N) = 0$$

Example: $I(N) = 1/N$
Second moment of number of objects affected in one timeslot = $o(N)$

- Similar result when mean field limit is in discrete time [Le Boudec et al 2007]
Example: Convergence to Mean Field

- Rescale time such that one time step = 1/N

- Number of transitions per time step is bounded by 2, therefore there is convergence to mean field

\[
\begin{align*}
\frac{\partial D}{\partial t} &= -\delta_D D - 2\lambda D^2 - \beta A \frac{D}{h+D} + (\alpha_0 + rD)S \\
\frac{\partial A}{\partial t} &= 2\lambda D^2 + \beta A \frac{D}{h+D} - \delta_A A + \alpha S \\
\frac{\partial S}{\partial t} &= \delta_D D + \delta_A A - (\alpha_0 + rD)S - \alpha S
\end{align*}
\]
Formulating the Mean Field Limit

- **Drift** = sum over all transitions of proba of transition $x$ Delta to system state $M^N(t)$

- Re-scale drift by intensity

- Equation for mean field limit is

  \[
  \frac{dm}{dt} = \text{limit of rescaled drift}
  \]

- Can be automated

  http://icawww1.epfl.ch/IS/tsed

<table>
<thead>
<tr>
<th>case</th>
<th>prob</th>
<th>effect on $(D, A, S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$D\delta_D$</td>
<td>$\frac{1}{N}(-1, 0, 1)$</td>
</tr>
<tr>
<td>2</td>
<td>$D\lambda N^{D-1} \frac{N-1}{N-1}$</td>
<td>$\frac{1}{N}(-2, +2, 0)$</td>
</tr>
<tr>
<td>3</td>
<td>$A\beta \frac{D}{h+D}$</td>
<td>$\frac{1}{N}(-1, +1, 0)$</td>
</tr>
<tr>
<td>4</td>
<td>$A\delta_A$</td>
<td>$\frac{1}{N}(0, -1, +1)$</td>
</tr>
<tr>
<td>5</td>
<td>$S(\alpha_0 + rD)$</td>
<td>$\frac{1}{N}(+1, 0, -1)$</td>
</tr>
<tr>
<td>6</td>
<td>$S\alpha$</td>
<td>$\frac{1}{N}(0, +1, -1)$</td>
</tr>
</tbody>
</table>

\[
drift = \left( -D\delta_D - 2D\lambda N^{D-1} \frac{N-1}{N-1} - A\beta \frac{D}{h+D} + S(\alpha_0 + rD) \right) \frac{1}{N} \\
2D\lambda N^{D-1} \frac{N-1}{N-1} + A\beta \frac{D}{h+D} - A\delta_A + S\alpha \\
D\delta_D + A\delta_A - S(\alpha_0 + rD) - S\alpha
\]

\[
\frac{\partial D}{\partial t} = -\delta DD - 2\lambda D^2 - \beta A \frac{D}{h+D} + (\alpha_0 + rD)S \\
\frac{\partial A}{\partial t} = 2\lambda D^2 + \beta A \frac{D}{h+D} - \delta AA + \alpha S \\
\frac{\partial S}{\partial t} = \delta DD + \delta AA - (\alpha_0 + rD)S - \alpha S
\]
Convergence to Mean Field

For the finite state space case, there are many simple results, often verifiable by inspection

For example [Kurtz 1970] or [Benaim, Le Boudec 2008]

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])
3. **FINITE HORIZON : FAST SIMULATION AND DECOUPLING ASSUMPTION**
Convergence to Mean Field Limit is Equivalent to Propagation of Chaos

Definition 1.1 Let $X^N = (X_1^N, ..., X_N^N)$ be an exchangeable sequence of processes in $\mathcal{P}(S)$ and $m \in \mathcal{P}(S)$ where $S$ is metric complete separable. $(X^N)_N$ is $m$-chaotic iff for every $k$: $\mathcal{L}(X_1^N, ..., X_k^N) \rightarrow m \otimes ... \otimes m$ as $N \rightarrow \infty$.

Theorem 1.1 ([Sznitman(1991)]) $(X^N)_N$ is $m$-chaotic then the occupancy measure $M^N \overset{\text{def}}{=} \frac{1}{N} \sum_{n=1}^{N} \delta_{X_n^N}$ converges in probability (and in law) to $m$.

If the occupancy measure converges in law to $m$ then $(X^N)_N$ is $m$-chaotic.
**Propagation of Chaos = Decoupling Assumption**

**Propagation of Chaos**

$k$ objects are asymptotically independent with common law equal to the mean field limit, for any fixed $k$

$$\mathcal{L} \left( X_1 \left( \frac{t}{l(N)} \right), ..., X_k \left( \frac{t}{l(N)} \right) \right) \rightarrow m(t) \otimes ... \otimes m(t)$$

**Decoupling Assumption**

*(also called Mean Field Approximation, or Fast Simulation)*

The law of one object is asymptotically as if all other objects were drawn randomly with replacement from $m(t)$
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 large $t$:
  - Prob (node $n$ is dormant) $\approx 0.3$
  - Prob (node $n$ is active) $\approx 0.6$
  - Prob (node $n$ is susceptible) $\approx 0.1$

- $m(t)$ approximates both
  1. the occupancy measure $M_N(t)$
  2. the state probability for one object at time $t$, drawn at random among $N$
\[ p_j^N(t|i) = P(X_n^N(t) = j \mid X_n^N(0) = i) \]
\[ p_j^N \left( \frac{t}{N} \mid i \right) \approx p_j(t|i) \]
where \( \hat{p}(t|i) \) is the (transient) probability of a continuous time nonhomogeneous Markov process
\[ \frac{d}{dt} \hat{p}(t|i) = \hat{p}(t|i)^T A(\bar{m}(t)) \]

- Same ODE as mean field limit, with different initial condition
\[ \frac{d}{dt} \bar{m}(t) = \bar{m}(t)^T A(\bar{m}(t)) = F(\bar{m}(t)) \]
The Decoupling Assumption

- 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 analyze or simulate evolution of \( k \) objects
4. INFINITE HORIZON: FIXED POINT METHOD AND DECOUPLING ASSUMPTION
The Fixed Point Method

- Decoupling assumption says distribution of prob for state of one object is $\approx \hat{m}(t)$ with $\frac{d\hat{m}(t)}{dt} = F(\hat{m}(t))$

- We are interested in stationary regime, i.e. we do $F(\hat{m}) = 0$

- This is the «Fixed Point Method»

- Example: in stationary regime:
  - Prob (node $n$ is dormant) $\approx 0.3$
  - Prob (node $n$ is active) $\approx 0.6$
  - Prob (node $n$ is susceptible) $\approx 0.1$

- Nodes $m$ and $n$ are independent
Example: 802.11 Analysis, Bianchi’s Formula

802.11 single cell
\[ m_i = \text{proba one node is in backoff stage I} \]

\[ \beta = \text{attempt rate} \]
\[ \gamma = \text{collision proba} \]

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

ODE for mean field limit

\[
\begin{align*}
\frac{dm_0}{d\tau} &= -m_0 q_0 + \beta (\bar{m}) (1 - \gamma (\bar{m})) + q_K m_K \gamma (\bar{m}) \\
\frac{dm_i}{d\tau} &= -m_i q_i + m_{i-1} q_{i-1} \gamma (\bar{m}) \quad i = 1, \ldots, K
\end{align*}
\]

\[ \beta (\bar{m}) = \sum_{i=0}^{K} q_i m_i \]
\[ \gamma (\bar{m}) = 1 - e^{-\beta (\bar{m})} \]

Solve for Fixed Point:

\[ m_i = \frac{\gamma^i}{q_i} \frac{1}{\sum_{k=0}^{K} \frac{\gamma^k}{q_k}} \]

Bianchi’s Fixed Point Equation

[Bianchi 1998]
2-Step Malware, Again

- Same as before except for one parameter value: $h = 0.1$ instead of 0.3

- The ODE does not converge to a unique attractor (limit cycle)

- The equation $F(m) = 0$ has a unique solution (red cross) – but it is not the stationary regime!
Example Where Fixed Point Method Fails

- 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’.

- This is synchronization.
Joint PDFs of Two Nodes in Stationary Regime

Stationary point of ODE

Mean of Limit of $\lim_{N \to \infty} \omega_N = \text{pdf of one node in stationary regime}$

pdf of node 2 in stationary regime, given node 1 is A

pdf of node 2 in stationary regime, given node 1 is S

pdf of node 2 in stationary regime, given node 1 is D
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.
Result 1: Fixed Point Method Holds under (H)

- Assume that

(H) ODE has a unique global stable point to which all trajectories converge

- Theorem [e.g. Benaim et al 2008] : The limit of stationary distribution of $M^N$ is concentrated on this fixed point

- The decoupling assumption holds in stationary regime
Result 2: Birkhoff Center

Here:
Birkhoff center = limit cycle $\cup$ fixed point

Theorem in [Benaim] says that the stochastic system for large $N$ is close to the Birkhoff center,

i.e. the stationary regime of ODE is a good approximation of the stationary regime of stochastic system
Stationary Behaviour of Mean Field Limit is not predicted by Structure of Markov Chain

- $M^N(t)$ is a Markov chain on $S^N=\{(a, b, c) \geq 0, a + b + c = 1, a, b, c$ multiples of $1/N\}$
- $M^N(t)$ is ergodic and aperiodic

Depending on parameter, there is or is not a limit cycle for $m(t)$

$h = 0.3$

$h = 0.1$
Example: 802.11 with Heterogeneous Nodes

[Cho et al, 2010]

Two classes of nodes with heterogeneous parameters (retransmission probability)

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

There is a limit cycle
Result 3: In the Reversible Case, the Fixed Point Method Always Works

■ **Definition** Markov Process $X(t)$ on enumerable state $E$ space, 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

1. Write dynamical system equations \textit{in transient regime}

2. Study the \textit{stationary regime of} dynamical system
   - \textbf{if} converges to unique stationary point $m^*$ \textbf{then} make fixed point assumption
   - \textbf{else} objects are coupled in stationary regime by mean field limit $m(t)$

Hard to predict outcome of 2 (except for reversible case)
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.

Stationary regime is more than stationary points, in general
(except for reversible case).
Thank You ...
References


[Tembine et al.(2009)] Tembine, Le Boudec, El-Azouzi, and Altman