Data-driven model learning in linear dynamic networks

Paul Van den Hof

ICMS Symposium, “Getting a grip on complex systems”
4 April 2022, Eindhoven

www.sysdynet.eu
www.pvandenhoef.nl
p.m.j.vandenhof@tue.nl
Introduction – dynamic networks

Decentralized process control

Smart power grid

Complex machines

Brain network

Hydrocarbon reservoirs

Pierre et al. (2012)

Mansoori (2014)

P. Hagmann et al. (2008)
Introduction

Overall trend:

- (Large-scale) interconnected dynamic systems
- Distributed / multi-agent type monitoring, control and optimization problems
- Data is “everywhere”, big data era, AI/machine learning tools
- Model-based operations require accurate/relevant models
- → Learning models from data (including physical insights when available)
Introduction

Drivers for data-processing / data-analytics:

Providing the tools for **online**
- Model estimation / calibration / adaptation

to accurately perform online model-based X:
- Monitoring
- Diagnosis and fault detection
- Control and optimization
- Predictive maintenance
- ......................
Introduction

Distributed / multi-agent control:

With both physical and communication links between systems $G_i$ and controllers $C_i$

How to address data-driven modelling problems in such a setting?
Introduction

The classical (multivariable) data-driven modeling problems\(^1\):

Identify a model of \( G \) on the basis of measured signals \( u, y \) (and possibly \( r \)), focusing on continuous LTI dynamics.

In interconnected systems (networks) the \textbf{structure / topology} becomes important to include.

\(^{[1]}\) Ljung (1999), Söderström and Stoica (1989), Pintelon and Schoukens (2012)
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification
• Global network identification
• Diffusively coupled networks
• Extensions - Discussion
Dynamic networks for data-driven modeling
Network models

D. Materassi and M.V. Salapaka (2012)
R.N. Mantegna (1999)
www.momo.cs.okayama-u.ac.jp
D. Koller and N. Friedman (2009)
J.C. Willems (2007)
E.A. Carara and F.G. Moraes (2008)
P.E. Paré et al (2013)
X. Cheng (2019)
Network models

\[
x(k + 1) = Ax(k) + Bu(k) \\
y(k) = Cx(k) + Du(k)
\]

- States as nodes in a (directed graph)
- State transitions (1 step in time) reflected by \( a_{i,j} \)
- Transitions are encoded in links
- Effect of transitions are summed in the nodes
- Self loops are allowed
- Actuation (\( u \)) and sensing (\( y \)) reflected by separate links

State space representation
Network models

\[ x(k + 1) = Ax(k) + Bu(k) \]
\[ y(k) = Cx(k) + Du(k) \]

- Ultimate break-down of structure in the system
- To smallest possible level of detail

For data-driven modeling problems:

- Stronger role for measurable inputs and outputs
- I/o dynamics can be lumped in dynamic **modules**
Network models

State space representation [1]

Module representation [2]

Dynamic network models - zooming

Decreasing structural information

Increasing level of detail
Dynamic network setup

- $v_i$: process noise
- $w_i$: node signal
- $r_i$: external excitation
- $G_{ij}$: module
Dynamic network setup

$G_{76}$ module
$r_i$ external excitation
$v_i$ process noise
$w_i$ node signal
Dynamic network setup

- $G_{76}$: module
- $r_i$: external excitation
- $v_i$: process noise
- $w_i$: node signal
Dynamic network setup

$G_{76}$ module

$r_i$ external excitation

$v_i$ process noise

$w_i$ node signal
Dynamic network setup

\[ r_i \quad \text{external excitation} \]

\[ v_i \quad \text{process noise} \]

\[ w_i \quad \text{node signal} \]
Dynamic network setup

Basic building block:

\[
w_j(t) = \sum_{k \in \mathcal{N}_j} G_{jk}^0(q) w_k(t) + r_j(t) + v_j(t)
\]

\(w_j\): node signal
\(r_j\): external excitation signal
\(v_j\): (unmeasured) disturbance, stationary stochastic process
\(G_{jk}^0\): module, rational proper transfer function, \(\mathcal{N}_j \subset \{\mathbb{Z} \cap [1, L]\}\backslash \{j\}\)
\(q\): shift operator, \(q^{-1}w(t) = w(t - 1)\)

Node signals: \(w_1, \cdots, w_L\)

Interconnection structure / topology of the network is encoded in \(\mathcal{N}_j, j = 1, \cdots, L\)
Dynamic network setup

Collecting all equations:

\[
\begin{bmatrix}
    w_1 \\
    w_2 \\
    \vdots \\
    w_L
\end{bmatrix} = \begin{bmatrix}
    0 & G^0_{12} & \cdots & G^0_{1L} \\
    G^0_{21} & 0 & \cdots & G^0_{2L} \\
    \vdots & \vdots & \ddots & \vdots \\
    G^0_{L1} & G^0_{L2} & \cdots & 0
\end{bmatrix} + R^0 \begin{bmatrix}
    r_1 \\
    r_2 \\
    \vdots \\
    r_K
\end{bmatrix} + H^0 \begin{bmatrix}
    e_1 \\
    e_2 \\
    \vdots \\
    e_p
\end{bmatrix}
\]

Network matrix $G^0(q)$

\[w(t) = G^0(q)w(t) + R^0(q)r(t) + v(t); \quad v(t) = H^0(q)e(t); \quad \text{cov}(e) = \Lambda\]

- Typically $R^0$ is just a (static) selection matrix, indicating which nodes have an excitation signal.
- The topology of the network is encoded in the structure (non-zero entries) of $G^0$.
- $r$ and $e$ are called external signals.
Dynamic network setup

Many challenging data-driven modeling questions can be formulated

Measured time series:
\[ \{ w_i(t) \}_{i=1,...,L}; \ \{ r_j(t) \}_{j=1,...,K} \]
Model learning problems

Under which conditions can we estimate the topology and/or dynamics of the full network?
How/when can we learn a local module from data (with known/unknown network topology)? Which signals to measure?
Model learning problems

Where to optimally locate sensors and actuators?
Model learning problems

Same questions for a subnetwork
Model learning problems

How can we benefit from known modules?
Model learning problems

Fault detection and diagnosis; detect/handle nonlinear elements
Model learning problems

Can we distribute the computations?
Dynamic network setup

- Identification of a local module (known topology)
- Identification of the full network
- Topology estimation
- Identifiability
- Sensor and excitation allocation
- Fault detection
- User prior knowledge of modules
- Distributed identification
- Scalable algorithms

Measured time series:
\[ \{w_i(t)\}_{i=1,...L}; \quad \{r_j(t)\}_{j=1,...K} \]
Dynamic network setup – directed graph

Nodes are vertices; modules/links are edges

Extended graph:
including the external signals
and disturbance correlations
Application: Networks of (damped) oscillators

- Power systems, vehicle platoons, thermal building dynamics, ...
- Spatially distributed
- Bilaterally coupled
Application: Printed Circuit Board (PCB) Testing

Detection of
• component failures
• parasitic effects

Source: Altium
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification
• Global network identification
• Diffusively coupled networks
• Extensions - Discussion
Single module identification
For a network with known topology:

• Identify $G_{21}^0$ on the basis of measured signals
• Which signals to measure? Preference for local measurements
• When is there enough excitation / data informativity?
Naïve approach: identify based on $w_1$ and $w_2$: in general does not work.
Identifying $G^0_{21}$ is part of a 4-input, 1-output problem

If noises $v_k$ are correlated it may even be part of a MIMO problem
Single module identification

Input signals will be correlated:
similar as in a closed-loop situation

What is required for
identifiability / data informativity?

Ability to distinguish between models
independent of id-method

Information content of signals
dependent on id-method

Identifiying $G^0_{21}$ is part of a
4-input, 1-output problem
Single module identification

Identifying $G_{21}^0$ is part of a 4-input, 1-output problem

All parallel paths, and loops around the output, plus input $w_1$ should have an independent external signal $r$ or $v$ and typically need to be blocked by a measured node

[1] Weerts et al., Automatica 2018, CDC 2018
[4] Shi et al., Automatica 2022
Single module identification

All inputs require an independent excitation (through vertex disjoint paths) from $r$, $e$

If excitation is relying on disturbances and correlated to $v_2$

To be handled by:
- Adding more input signals (blocking the cv)
- Including the input as output (MIMO) [3]

Confounding variable [1][2]

Single module identification

Typical solution:

- MISO (sometimes MIMO) estimation problem
- to be solved by any (closed-loop) identification algorithm, e.g. direct/indirect method
Machine learning in local module identification

- MISO identification with all modules parameterized
- Brings in two major problems:
  - Large number of parameters to estimate
  - Model order selection step for each module (CV, AIC, BIC)

- For 5 modules, combinations = 244,140,625

  Increases variance
  Computationally challenging

- We need only the target module. No NUISANCE!
Maximize marginal likelihood of output data: $\hat{\eta} = \arg\max \ p(w_j; \eta)$

$\eta := [\theta \ \lambda_j \ \lambda_{k_1} \ \ldots \ \lambda_{k_p} \ \beta_j \ \beta_{k_1} \ \ldots \ \beta_{k_p} \ \sigma_j^2]^T$

- smaller no. of parameters
- simpler model order selection step
- scalable to large dynamic networks
- simpler optimization problems to estimate parameters

Numerical simulation

- Identify $G_{31}$ given data
- 50 independent MC simulation
- Data = 500
Summary single module identification

• Path-based conditions for **network identifiability** (where to excite?)

• Graph tools for checking conditions

• Degrees of freedom in selection of measured signals – sensor selection

• Methods for **consistent** and **minimum variance** module estimation, and effective (scalable) algorithms

• A priori known modules can be accounted for
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification
• Global network identification
• Diffusively coupled networks
• Extensions - Discussion
Under which conditions can we estimate the topology and/or dynamics of the full network?
Network identifiability

Question: Can different dynamic networks be distinguished from each other from measured signals $w$, $r$?
Network identifiability

The identifiability problem:

The network model:

\[ w(t) = G(q)w(t) + R(q)r(t) + \underbrace{H(q)e(t)}_{v(t)} \]

can be transformed with any rational \( P(q) \):

\[ P(q)w(t) = P(q)\{G(q)w(t)+R(q)r(t)+H(q)e(t)\} \]

to an equivalent model:

\[ w(t) = \tilde{G}(q)w(t) + \tilde{R}(q)r(t) + \tilde{H}(q)e(t) \]

**Nonuniqueness**, unless there are structural constraints on \( G, R, H \).

[2] Bottegal et al., SYSID 2017
Network identifiability

Consider a network model set:

$$\mathcal{M} = \{(G(\theta), R(\theta), H(\theta))\}_{\theta \in \Theta}$$

representing structural constraints on the considered models:

- modules that are fixed and/or zero (topology)
- locations of excitation signals
- disturbance correlation

Generic identifiability of $\mathcal{M}$:

- There do not exist distinct equivalent models (generating the same data)
- for almost all models in the set.

[1] Weerts et al., SYSID2015; Weerts et al., Automatica, March 2018;
Example 5-node network

Conditions for identifiability rank conditions on transfer function

Full row rank of

\[
\begin{bmatrix}
  v_3 \\
r_4 \\
r_5
\end{bmatrix} \rightarrow \begin{bmatrix}
w_2 \\
w_5
\end{bmatrix}
\]

For the **generic case**, the rank can be calculated by a graph-based condition\(^1,\)^2:

**Generic rank = number of vertex-disjoint paths**

2 vertex-disjoint paths $\rightarrow$ full row rank 2

The rank condition has to be checked for all nodes.

---

\(^1\) Van der Woude, 1991
\(^2\) Hendrickx, Gevers & Bazanella, CDC 2017, TAC 2019
Synthesis solution for network identifiability

Allocating external signals for generic identifiability:

1. Cover the graph of the network model set by a set of disjoint pseudo-trees
   Pseudo-trees:
   Tree with root in green
   Cycle with outgoing trees; Any node in cycle is root
   Edges are disjoint and all out-neighbours of a node are in the same pseudo-tree

2. Assign an independent external signal (\( r \) or \( e \)) at a root of each pseudo-tree.
   This guarantees generic identifiability of the model set.

Where to allocate external excitations for network identifiability?

All indicated modules are parametrized

Two disjoint pseudo-trees
Where to allocate external excitations for network identifiability?

Two independent excitations guarantee generic network identifiability

Where to allocate external excitations for network identifiability?

- Nodes are signals $w$ and external signals $(r, e)$ that are input to parametrized link
- Known (nonparametrized) links do not need to be covered

---

Summary identifiability of full network

Identifiability of network model sets is determined by

- Presence and location of external signals, and
- Correlation of disturbances
- Topology of parametrized modules

- Graphic-based tool for synthesizing allocation of external signals

Extensions:
- Situations where not all node signals are measured \(^1\)

\(^1\) Bazanella, CDC 2019.
Algorithms for identification of full network

(Prediction error) identification methods will typically lead to large-scale non-convex optimization problems

Convex relaxation algorithms are being developed\textsuperscript{[1,2]} as well as machine learning tools

\textsuperscript{[1]} Weerts, Galrinho et al., SYSID 2018
\textsuperscript{[2]} Fonken et al., Automatica 2022, to appear.
Topology identification

• Topology resulting from full dynamic model

• Alternative: non-parametric models (Wiener filters\[^1\]) or kernel-based approaches\[^2\][\[^3\]]

• Modeling module dynamics by Gaussian processes,
  
  kernel with 2 parameters for each dynamic module

• Optimizing likelihood of the data as function of parameters and topology:

\[ p(\{w(t)\}_{t=1}^{N}|\theta, \mathcal{G}) \]

• Forward-backward search over topologies + empirical Bayes (EM) for parameters

[3] Shi, Bottegal, PVdH, ECC 2019
Topology identification

50 MC realizations of network with 6 nodes.

[1] Shi, Bottegal, PVdH, ECC 2019
Neurodynamic effect of listening to Mozart music

Identifying changes in network connections in the brain, after intensely listening for one week

Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification
• Global network identification
• Diffusively coupled networks
• Extensions - Discussion
Diffusively coupled networks
Back to the basics of physical interconnections

In connecting physical systems, there is often no predetermined direction of information \[1\]

Example: resistor / spring connection in electrical / mechanical system:

\[
I = \frac{1}{R} (V_1 - V_2)
\]

\[
F = K (x_1 - x_2)
\]

Difference of node signals drives the interaction: **diffusive coupling**

Diffusively coupled physical network

Equation for node $j$:

$$M_j \ddot{w}_j(t) + D_{j0} \dot{w}_j(t) + \sum_{k \neq j} D_{jk}(\dot{w}_j(t) - \dot{w}_k(t)) + K_{j0} w_j(t) + \sum_{k \neq j} K_{jk}(w_j(t) - w_k(t)) = u_j(t),$$
Mass-spring-damper system

- Masses $M_j$
- Springs $K_{jk}$
- Dampers $D_{jk}$
- Input $u_j$

\[
\begin{bmatrix}
  M_1 & M_2 & M_3 \\
  M_2 & M_3 & M_3 \\
  M_3 & M_3 & M_3
\end{bmatrix}
\begin{bmatrix}
  \ddot{w}_1 \\
  \ddot{w}_2 \\
  \ddot{w}_3
\end{bmatrix} + \begin{bmatrix}
  0 & D_{20} & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
  \dot{w}_1 \\
  \dot{w}_2 \\
  \dot{w}_3
\end{bmatrix} + \begin{bmatrix}
  0 & K_{10} & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
  w_1 \\
  w_2 \\
  w_3
\end{bmatrix} \\
+ \begin{bmatrix}
  D_{13} & 0 & -D_{13} \\
  0 & D_{23} & -D_{23} \\
  -D_{13} & -D_{23} & D_{13} + D_{23}
\end{bmatrix}
\begin{bmatrix}
  \ddot{w}_1 \\
  \ddot{w}_2 \\
  \ddot{w}_3
\end{bmatrix} + \begin{bmatrix}
  K_{12} + K_{13} & -K_{12} & -K_{13} \\
  -K_{12} & K_{12} & 0 \\
  -K_{13} & 0 & K_{13}
\end{bmatrix}
\begin{bmatrix}
  w_1 \\
  w_2 \\
  w_3
\end{bmatrix} = \begin{bmatrix}
  0 \\
  u_2 \\
  0
\end{bmatrix}
\]

\[
\begin{bmatrix}
  A(p) \\
  B(p)
\end{bmatrix}_{diagonal, \text{Laplacian}} w(t) = u(t) \quad A(p), B(p) \text{ polynomial} \quad p = \frac{d}{dt}
\]
Mass-spring-damper system

\[
\begin{bmatrix}
A(p) + B(p)
\end{bmatrix} w(t) = u(t) \quad A(p), B(p) \text{ polynomial}
\]

\[
\begin{bmatrix}
Q(p) - P(p)
\end{bmatrix} w(t) = u(t)
\]

This fully fits in the earlier module representation:

\[
w(t) = Gw(t) + \underbrace{Rr(t) + He(t)}_{Q^{-1}(p)u(t)}
\]

with the additional condition that:

\[G(p) = Q(p)^{-1}P(p)\] \quad \text{\(Q(p), P(p)\) polynomial}

\[P(p)\text{ symmetric, } Q(p)\text{ diagonal}\]
Module representation

Consequences for node interactions:

- Node interactions come in pairs of modules
- Where numerators are the same

Framework for network identification remains the same

- Symmetry can simply be incorporated in identification
Local network identification

Identification of **one** (physical) interconnection

Identification of **two** modules $G_{jk}$ and $G_{kj}$
Immersion conditions

For simultaneously identifying two modules in one interconnection:

The parallel path and loops-around-the-output condition, now simplifies to:

Measuring/exciting all neighbouring nodes of \( w_2 \) and \( w_3 \) leads to a solution

E.E.M. Kivits et al., CDC 2019.
Summary diffusively coupled networks

• Diffusively coupled networks fit within the module framework (special case)
  - no restriction to second order equations

• Earlier identification framework can be utilized

• Local identification is well-addressed (and stays really local)

• Framework is fit for representing cyber-physical systems
  (combining physical bi-directional links, and cyber uni-directional links).
Extensions - Discussion
Extensions - Discussion

• Including sensor noise \[^{[1]}\]
  • Errors-in-variabeals problems can be more easily handled in a network setting

• Distributed estimation (MISO models) \[^{[2]}\]
  • Communication constraints between different agents
  • Recursive (distributed) estimator converges to global optimizer (more slowly)

• Experiment design \[^{[3],[4]}\]
  • design of least costly experiments

Summary

• **Dynamic network modeling:**
  intriguing research topic with many open questions
• The (centralized) LTI framework is only just the beginning
• Further move towards data-aspects related to distributed control
• and large-scale aspects
• and more real-life applications
Matlab Toolbox
ERC SYSDYNET Team: data-driven modeling in dynamic networks

Research team:

Karthik Ramaswamy  
Arne Dankers  
Harm Weerts  
Shengling Shi  
Giulio Bottegal  
Xiaodong Cheng  
Mannes Dreef  
Lizan Kivits  
Tom Steentjes  
Stefanie Fonken  
Mircea Lazar  
Tijs Donkers  
Jobert Ludlage

Co-authors, contributors and discussion partners:
Donatello Materassi, Manfred Deistler, Michel Gevers, Jonas Linder, Sean Warnick, Alessandro Chiuso, Håkan Hjalmarsson, Miguel Galrinho, Martin Enqvist, Johan Schoukens, Xavier Bombois, Peter Heuberger, Péter Csurcsia
Minneapolis, Vienna, Louvain-la-Neuve, Linkoping, KTH Stockholm, Padova, Brussels, Salt Lake City, Lyon.
Further reading


Papers available at www.pvandenhof.nl
The end