**INTRODUCTION/ABSTRACT**

In many problems it is helpful and even necessary to study the relationships between entities of a system since we find them interconnected building a network in which some small changes on an element can propagate and affect the entire system. We find a large list of examples in which it is necessary to apply network theory on different and wide subjects as social science, finance, biology or climatology.

Studying such networks can reveal lots of information like how an event can trigger a topological change of the entire network, how entities of the network can depend on each other or share similar properties and organize themselves into groups or which individuals are the main agents in the network ruling its evolution.

Classically these analysis have been carried in the simplest way assuming they are static however we truly know that this is false in most of the cases. The internet, biological processes in a body or the economy are some examples in which networks change and evolve over time shrinking or creating new links as the attributes or states of the individuals change.

Although there is wide literature and methods on modelling static networks (REFERENCIAS) it has not been until recent years that interest has grown on studying dynamic or time-varying networks.

A **time-varying network**, also known as a temporal network, is a [network](https://en.wikipedia.org/wiki/Complex_network) whose links are active only at certain points in time. Each link carries information on when it is active, along with other possible characteristics such as a [weight](https://en.wikipedia.org/wiki/Weighted_networks). Time-varying networks are of particular relevance to [spreading processes](https://en.wikipedia.org/wiki/Complex_contagion), like the spread of information and disease, since each link is a contact opportunity and the time ordering of contacts is included.

Time-varying networks are characterized by intermittent activation at the scale of individual links. This is in contrast to various models of [network evolution](https://en.wikipedia.org/wiki/Evolving_networks), which may include an overall time dependence at the scale of the network as a whole.

The study of these networks can be done from different points of view, depending on the problem under consideration it can be convenient to study the evolution of the relationships (edge-centric), one or more entities (vertex-centric) or the global system (graph-centric).

Usually, only time series measurements, such as microarray, stock price, etc., of the activity of the nodal entities, but not their linkage status, are available. The goal is to recover the latent time-varying networks with temporal resolution up to every single time point based on time series measurements.

To do so this masters’ project is going to focus on a Machine Learning algorithm called TESLA, built to recover the structure of time-varying networks over a fixed set of nodes from these time series nodal attributes (It is obvious that in these networks an edge or link between two nodes implies dependency between their nodal states), aiming to recover finance network structures.

To be able to understand the current state of the research on the field and the strengths, weakness and challenges of the techniques proposed some literature review has been carried out.

**CURRENT MAIN ALGORITHMS**

A classical model of static network analysis is the exponential random graph model (ERGM) (Guo F, Hanneke S, Fu W, Xing E (2007)) that was used as a base for two of the first algorithms proposed to study dynamic networks, the temporal exponential random graph model (tERGM) and the hidden tERGM for modelling a sequence of node attribute observations. When tERGM assume that the sequence of networks is available, htERGM explores the possible dependencies of unobserved rewiring networks and leads to the algorithm that can reconstruct such networks from a snapshots’ sequence of nodal attributes.

Although this last algorithm overcame approaches that recover single time-invariant networks it is not the most useful one since it depends on unobserved network variables being unable to compute likelihood ratios, needing inference algorithms specifically for each problem.

Another classical approach is using Dynamic Bayesian Networks (DBN), that have been widely used to recover gene regulatory relationships from time-series data in computational systems biology. Its standard assumption is ‘stationarity’, and therefore, several research efforts have been recently proposed to relax this restriction. However, those methods suffer from three challenges: long running time, low accuracy and reliance on parameter settings. [Zou & Wang 2015] propose a novel non-stationary DBN model by extending each hidden node of Hidden Markov Model into a DBN (called HMDBN), which properly handles the underlying time-evolving networks resulting in a promising experimental evaluation of the method, demonstrating more stably high prediction accuracy and significantly improved computation efficiency (even with no prior knowledge and parameter settings) on both synthetic and real biological data.

Although this probabilistic model is a lot more complex than the TESLA algorithm it could be a good method to implement in order to compare results and performance.

In this line Durante & Dunson 2014 proposed a Bayesian non-parametric model including time-varying predictors in dynamic network inference precisely for financial studies.

This model computes edge specific predictors where the link probabilities () are estimated via a logistic regression, with a baseline process quantifying the overall propensity to form links in the network across time, are vectors containing the latent coordinates favouring a higher link probability when units *i* and *j* have latent coordinates in the same direction and is a P-dimensional vector of time-varying edge-specific predictors for units *i* and *j* at time t and are the corresponding dynamic coefficients. This allows the proximity between units *i* and *j* at time *t* to depend on predictors in a manner that varies smoothly with time.

The main issue regarding this method is that it assumes time-constant smoothness while in finance and other network applications we expect smoothness to vary over time.

TESLA, from the acronym TESLLOR, which stands for temporally smoothed *l*1*-*regularized logistic regression represents an extension of the lasso-style sparse structure recovery technique and is based on a key assumption that temporally adjacent networks are likely not to be dramatically different from each other in topology and therefore are more likely to share common edges than temporally distant networks.

Building on the *l*1-regularized logistic regression algorithm for estimating single sparse networks (*Wainwright M, Ravikumar P, Lafferty J (2006) High dimensional graphical model selection using l1-regularized logistic regression. Advances in Neural Information Processing Systems 19 (MIT Press, Cambridge, MA), pp 1465–1472*) it was developed a regression regularization scheme that connects multiple time-specific network inference functions via a first-order edge smoothness function that encourages edge retention between time-adjacent networks. An important property of this idea is that it fully integrates all available samples of the entire time series in a single inference procedure, what means an advantage in contrast of htERGM algorithm.

TESLA estimates , the correlation (or dependency strength) matrix between the nodes using a time-series of the observed stated of the nodes **xt**.

Where

Where the graph structure is given by the locations of the nonzero elements of the parameter vectors . Components of the vectors are indexed by distinct pairs of nodes and a component *j* of the vector is nonzero if and only if the corresponding edge (*i, j*) ∈ Eτ. denotes the observed states of all nodes but node *i* and is the log conditional likelihood of state under a logistic regression model.

First term of the algorithm are p logistic regressions, one for each node with respect the rest of them ( are the states of all nodes but node I in the dth sample in time epoch t).

The second term is a lasso regularization that introduces sparsity and consistency in neighbourhood selection, as seen in (*Wainwright M, Ravikumar P, Lafferty J (2006) High dimensional graphical model selection using l1-regularized logistic regression. Advances in Neural Information Processing Systems 19 (MIT Press, Cambridge, MA), pp 1465–1472.*). Wainwright et al. have shown that that applying this l1 norm-regularized logistic regression leads to consistent neighbourhood selection, so it converges to the true graph structure. Also a sparse graph effectively limits the degree of freedom of the model, which makes structure recovery possible given a small sample size (*Kolar M., Song L, Ahmed A., Xing E. P. Estimating time-varying networks. Institute of Mathematical Statistics in The Annals of Applied Statistics, 2010, Vol. 4, No. 1, 94–123*)

Since in this problem we are estimating dynamic networks (so more than one single graph), and, as assumed before, temporary adjacent networks are going to be similar one to each other, we need to introduce the third term in the equation that penalises the discrepancy between time-adjacent parameters .

Summarizing, the first term is the main one while the last two terms are the penalties () that are introduced to enforce sparsity and smoothness.

It is important to note that in this specific method we are assuming that are a piecewise constant function with abrupt changes in parameters (jumps).

However in some cases we need the changes in parameters to be even less abrupt, so further work adapted this algorithm with a weighting term that works with as real smooth functions instead of a penalty in the structural changes () that assumes “jumps” between adjacent values.

Where the weights are defined by:

And is a symmetric nonnegative kernel function.

FIT ALL IN THE PAPER-TEMPLATE

SPECIFIC PROBLEM TO STUDY: FINANCE ???

The problem of dynamic structure estimation is of high importance in domains that lack prior knowledge or measurement techniques about the interactions between different actors; and such estimates can provide .desirable information about the details of relational changes in a complex system

Consider the following problem:

Understanding stock market. In a finance setting we have values of different stocks at each time point. Suppose, for simplicity, that we only measure whether the value of a particular stock is going up or down. We would like to find the underlying transient relational patterns between different stocks from these measurements and get insight into how these patterns change over time.

A key technical hurdle preventing us from an in-depth investigation of the mechanisms underlying these complex systems is the unavailability of serial snapshots of the time-varying networks underlying these systems

*Kolar M., Song L, Ahmed A., Xing E. P. Estimating time-varying networks. Institute of Mathematical Statistics in The Annals of Applied Statistics, 2010, Vol. 4, No. 1, 94–123*

Whether using time-varying networks will be worth the added complexity depends on the relative [time scales](https://en.wikipedia.org/wiki/Time_scales) in question

Time-varying networks are most useful in describing systems where the spreading process on a network and the network itself evolve at similar timescales

In time-varying networks the timescale for the evolution of the network as a whole is comparable to the timescale of the evolution of the spreading process, so the interplay between them becomes important (*P. Holme, J. Saramäki. Temporal Networks. Phys. Rep. 519, 99–100; 10.1016/j.physrep.2012.03.001 (2012))*

The most challenging aspect in estimating time-varying graphs is that the dimension of the data p can be much larger than the size of the sample n (p ≫ n), and there is usually only one sample per time point. Ejemplo finanzas: p number of stocks (=100?) n …. . Then, the question is what are the sufficient conditions under which our algorithms recover the sequence of unknown graphs correctly

It might seem that the problem is ill-defined, since for any time point we have at most one observation; however, as we will show shortly, under a set of suitable assumptions the problem is indeed well defined and the series of underlying graph structures can be estimated.

Another popular approach to the graph structure estimation is the ℓ1 penalized likelihood maximization, which simultaneously estimates the graph structure and the elements of the covariance matrix, however, at a price of computational efficiency. The penalized likelihood approach involves solving a semidefinite program (SDP) and a number of authors have worked on efficient solvers that exploit the special structure of the problem [Banerjee, Ghaoui and d’Aspremont (2008); Yuan and Lin (2007); Friedman, Hastie and Tibshirani (2007); Duchi, Gould and Koller (2008); Rothman et al. (2008)]. Of these methods, it seems that the graphical lasso [Friedman, Hastie and Tibshirani (2007)] is the most computationally efficient. Some authors have proposed to use a nonconcave penalty instead of the ℓ1 penalty, which tries ESTIMATING TIME-VARYING NETWORKS 5 to remedy the bias that the ℓ1 penalty introduces [Fan and Li (2001); Fan, Feng and Wu (2009); Zou and Li (2008)].

CONCLUSION

Summarizing we find interesting studying data through a network lens since it can add substantial new insights. This work evaluated the state-of-the-art methods in the dynamic networks field in order to prepare further research of applying (mainly) TESLA algorithm in a finance context to study the changes that may warn or even drive significant global events like the 2008 crisis.

It is important to use scalable computational methods that can fit very large networks (huge number of nodes), so models like TESTLA that exploit sparsity are promising in this matter.

Reviewing the two TESLA-based algorithms this project is going to be focused on we find they represent two different ends: one is tailored toward estimation of structural changes in the model while the other is able to estimate smoothly changing networks. This gives us a great opportunity to study the same field with these two different approaches and test which one better models the stocks’ relationships.

**REFERENCES**

Guo F, Hanneke S, Fu W and Xing E (2007) Recovering temporally rewiring networks: A model-based approach. *The 24th International Conference of Machine Learning*, *2007* (Association for Computing Machinery, New York).

Zhu S. and Wang Y. (2015). Hidden Markov induced Dynamic Bayesian Network for recovering time evolving gene regulatory networks. *Scientific Reports*, *5*, 17841. <http://doi.org/10.1038/srep17841>

Durante D. and Dunson D. (2014) Bayesian dynamic financial networks with time-varying predictors. *Statistics & Probability Letters*, 93, 19-26.

Ahmed A and Xing E. Recovering time-varying networks of dependencies in social and biological stud**i**es. *PNAS* (2009) 106 (29) 11878-11883.

Kolar M, Song L, Ahmed A and Xing, Eric P. Estimating time-varying networks. *Annals Applied Statistics Vol. 4* (2010). No. 1, 94-123.

Añadir las subreferencias!!!