Age-Dependent Analysis of Mortality Patterns in Italy: A Network Perspective via Dynamic Stochastic Block Models

Cristian Castiglione[1(✉)] and Giovanni Romanò[2]

[1] Bocconi Institute for Data Science and Analytics, Bocconi University, Milan, Italy
cristian.castiglione@unibocconi.it
[2] Department of Decision Sciences, Bocconi University, Milan, Italy
giovanni.romano4@phd.unibocconi.it

Abstract. Discovering latent dependence structures over the nodes of a dynamic network is a difficult challenge that is of increasing importance in many applied fields. Our interest is motivated by a demographic analysis of the complex interaction between causes of death in the Italian population observed on a fine age grid. To unveil non-trivial grouping structures between causes of death, we rely on a simplified version of a recently proposed stochastic block model for dynamic networks [4], which is able to learn node partitions having common connectivity patterns. To flexibly account for the time evolution of the node grouping structure, such a model relies on a dynamic random partition process, which permits to learn sequences of partitions with a high and evolving level of persistence over time.

Keywords: Causes of death data · Dependent partitions · Sequences of networks · Stochastic block models

1 Introduction

Network analysis has become a central tool in statistics and machine learning for discovering non-trivial dependence structures in complex systems characterized by the interaction of several units, called *nodes*, whose relationships can be described as the *edges* of a graph. Relevant applications include genomics, epidemiology, and social sciences. In particular, a key challenge in network analysis is community detection, which uncovers structural grouping patterns within interconnected data.

This study is driven by the demographic challenge of analyzing the age-dependent evolution of the interactions among causes of death in the Italian population. Understanding the distribution, evolution, and structural dependence of mortality causes is crucial for demographic and public health planning. While previous studies have ranked the leading causes of death and examined their trends, they often overlook the complexity of interactions among finer categorizations of causes [1,5]. Only recently [4] addressed this gap by taking a network

E. di Bella et al. (Eds.): SIS 2025, ISSSAS, pp. 271–276, 2025.
https://doi.org/10.1007/978-3-031-96736-8_45

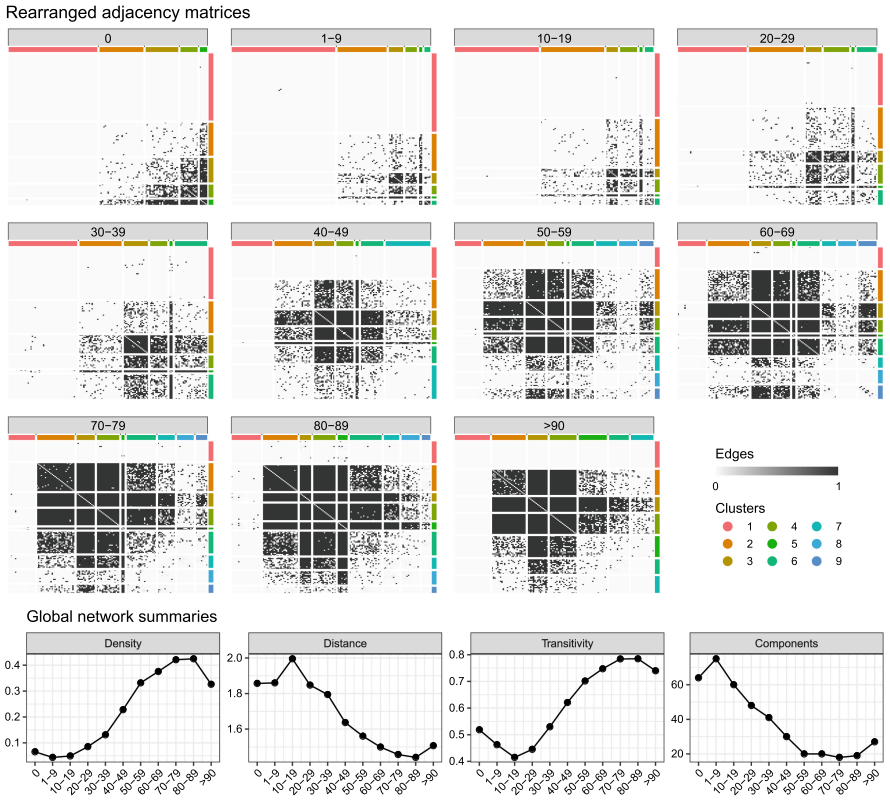

Fig. 1. Causes-of-death network sequence for the male population recorded in Italy in 2019. Top: adjacency matrices, where black and white cells represent edges and nonedges, respectively. Colored strips identify the partition estimated using dSBM (see Sects. 2 and 3). Bottom: network summary indicators over age.

viewpoint and representing the causes of death as the nodes of an undirected graph, where edges encode age-specific co-occurrences of different mortality factors. Figure 1 shows the adjacency matrices and some global summary statistics of the networks obtained for the male population in Italy recorded in 2019. The considered networks contain 139 nodes each derived by the ICD-10 classification [7], where we excluded external causes, injuries, and never-observed causes. Each network is built on an evenly spaced age grid covering 10 years of collected data, while the full sequence of networks spans all available ages. Deaths occurring within the first year of life are treated separately to account for delivery and infant mortality. An edge is observed in the network if at least 1 co-occurrence has been observed in the population during the considered age period. Both the adjacency matrices and the summary statistics manifest a clear dynamic evolution of the network sequence during the life span, where at older ages the networks exhibit increasingly higher connectivity and transitivity. The raw data are

available upon request from the Italian National Institute of Statistics (ISTAT), which collects all the mortality causes for each recorded death.

Assuming that the mortality causes have a mutually exclusive and exhaustive grouping structure, with nodes in the same cluster sharing common connectivity profiles at a fixed age, we employ a simplified version of the *dynamic Stochastic Block Model* (dSBM) proposed in [4] to flexibly characterize the network evolution over age. Such a simplified version leverages an *Extended Stochastic Block Model* (ESBM) [2] likelihood and a *time-dependent Random Partition Model* (tRPM) [3] prior to infer hidden structures in the network sequence, allowing for dynamic clustering of causes over age or time. This method overcomes key limitations of existing statistical models for dynamic networks by jointly estimating the number of clusters and modeling temporal dependencies.

2 Dynamic Stochastic Block Model

Here, we briefly outline a simplified version of the dSBM model proposed by [4], focusing on the likelihood specification in Sect. 2.1, prior elicitation in Sect. 2.2, and posterior computation in Sect. 2.3. Unlike the general dSBM model, such a simplified version assumes binary undirected edges, and a single partition sequence regulating the node group structure. By doing so, we first introduce the following notation. Let Y_x denote the $V \times V$ symmetric adjacency matrix at age $x \in \{1, \dots, N\}$ encoding the binary relationships between nodes, i.e., causes of death. In particular, each $y_{vux} = y_{uvx} \in \{0, 1\}$ account for the symmetric link between v and u, for every $v \in \{2, \dots, V\}$ and $u \in \{1, \dots, v-1\}$, with y_{vux} taking value 1 if an edge is observed and 0 otherwise.

2.1 Likelihood Specification

Specializing the general dSBM to our data, we assume for the causes-of-death network at age x an ESBM specification [2], clustering the nodes into mutually exclusive and exhaustive groups. In particular, we denote by $z_x = (z_{1x}, \dots, z_{Vx})$ the vector of cluster allocations, where $z_{vx} = h \in \{1, \dots, H_x\}$ means that at age x the cause v belongs to group h. Additionally, we define $\rho_x \in P$ as the partition induced by the allocation vector z_x, where P is the set of all possible partitions of V nodes. dSBM allows for a different number of groups at each age, letting H_x vary over x. To accommodate for the dichotomic nature of the adjacency matrix, we consider a Bernoulli distribution with parameter $\theta_{z_{vx}z_{ux}x} \in (0, 1)$ for y_{vux}. Moreover, we assume homogeneous connectivity structures within and between each block, that is $\theta_{z_{vx}z_{ux}x} = \theta_{hkx}$ for $z_{vx} = h$ and $z_{ux} = k$. Finally, we consider an independent Beta prior distribution on each θ_{hkx}, thus leading to the hierarchical model

$$(y_{vux} \mid z_{vx} = h, z_{ux} = k, \theta_{hkx}) \overset{\text{ind}}{\sim} \text{Bernoulli}(\theta_{hkx}), \quad \theta_{hkx} \overset{\text{ind}}{\sim} \text{Beta}(a_\theta, b_\theta),$$

for nodes $v \in \{2, \dots, V\}$ and $u \in \{1, \dots, v-1\}$, cluster labels $h, k \in \{1, \dots, H_x\}$ and age class $x \in \{1, \dots N\}$, where $a_\theta, b_\theta > 0$ are fixed prior hyperparameters.

2.2 Prior Specification

To model the age dependence between subsequent networks, dSBM relies on a latent partition process following a Markovian evolution of the partitions, that is $\text{pr}(\rho_1, \dots, \rho_N) = \text{pr}(\rho_1) \prod_{x=2}^{N} \text{pr}(\rho_x \mid \rho_{x-1})$. Then, we need to specify an explicit form for the transition probability $\text{pr}(\rho_x \mid \rho_{x-1})$ and the initial distribution $\text{pr}(\rho_1)$. To this end, dSBM employs the tRPM prior proposed by [3], which specifies the joint distribution of the partition sequence by introducing the augmented node-specific variables $\gamma_x = (\gamma_{2x}, \dots, \gamma_{Vx})$, such that

$$\text{pr}(\rho_1, \gamma_2, \rho_2, \dots, \gamma_N, \rho_N) = \text{pr}(\rho_1) \prod_{x=2}^{N} \text{pr}(\gamma_x) \, \text{pr}(\rho_x \mid \gamma_x, \rho_{x-1}),$$

where $\text{pr}(\rho_1)$ is an exchangeable partition probability function (EPPF) describing how V nodes are clustered in H_1 distinct groups, and γ_x is a parameter vector controlling the similarity between ρ_x and ρ_{x-1}. In particular, auxiliary variables $\gamma_x \in \{0,1\}^V$ are node-specific indicators that identify which nodes can be considered for cluster reallocation moving from age $x-1$ to age x, that is: $\gamma_{vx} = 0$ if node v can be reallocated, $\gamma_{vx} = 1$ otherwise. tRPM then assumes independent Bernoulli distributions with parameter $\alpha_x \in [0,1]$ for all auxiliary variables, such that α_x plays the role of a common dependence parameter. Indeed, $\alpha_x = 0$ implies almost sure independence between ρ_x and ρ_{x-1}, conversely, $\alpha_x = 1$ implies that $\rho_x = \rho_{x-1}$ with probability 1. In addition, tRPM assumes Beta prior distributions over α_x, thus obtaining the hierarchical prior specification

$$(\gamma_{vx} \mid \alpha_x) \overset{\text{ind}}{\sim} \text{Bernoulli}(\alpha_x), \quad \alpha_x \overset{\text{ind}}{\sim} \text{Beta}(a_\alpha, b_\alpha),$$

for node $v \in \{1, \dots, V\}$ and age $x \in \{1, \dots N\}$, where $a_\alpha, b_\alpha > 0$ are fixed prior hyperparameters.

To characterize the transition probability $\text{pr}(\rho_x \mid \gamma_x, \rho_{x-1})$, we define the index set $R_x = \{v : \gamma_{vx} = 1\}$ as the collection of nodes that can not be reallocated. Moreover, we define $\rho_x^{R_x}$ as the reduced partition corresponding to the index set R_x. Then, defining the restricted partition space $P_x = \{\rho_x \in P : \rho_x^{R_x} = \rho_{x-1}^{R_x}\}$, tRPM assumes for the transition probability from ρ_{x-1} to ρ_x the following form

$$\text{pr}(\rho_x = \rho \mid \gamma_x, \rho_{x-1}) \propto \text{pr}(\rho_x = \rho) \, \mathbb{1}(\rho \in P_x),$$

where $\text{pr}(\rho_x = \rho)$ is the EPPF at age 1 evaluated at the generic partition $\rho \in P$, and $\mathbb{1}(A)$ is the indicator function of the event A.

To complete the prior specification for the partition sequence $\{\rho_1, \dots, \rho_N\}$, tRPM considers a Chinese Restaurant Process prior with concentration parameter $\eta > 0$ for the initial partition ρ_1, thus implying the following marginal EPPF for ρ_x: $\text{pr}(\rho_x = \rho) = \eta^{H_x} \Gamma(\eta) \prod_{h=1}^{H_x} \Gamma(V_{hx}) / \Gamma(\eta + V)$, for all $x \in \{1, \dots, N\}$, where V_{hx} is the size of cluster h at age x. Additionally, we assume the conjugate prior distribution $\eta \sim \text{Gamma}(a_\eta, b_\eta)$, where $a_\eta, b_\eta > 0$ are fixed prior hyperparameters, enabling a flexible estimation of the concentration parameter.

2.3 Posterior Computation

Similarly to the general dSBM in [4], also our simplified version outlined in Sects. 2.1 and 2.2 does not enjoy a posterior distribution of known closed-form, hence MCMC algorithms are needed to perform posterior inference. To this end, we adapt the collapsed Gibbs sampler in [4] to our construction. The resulting routine iteratively samples from the full-conditional distributions $(\gamma \mid z, \alpha, \eta, y)$, $(z \mid \gamma, \alpha, \eta, y)$, $(\alpha \mid z, \gamma, \eta, y)$ and $(\eta \mid z, \gamma, \alpha, y)$, where the edge probabilities θ can be marginalized analytically, as shown in [2]. To summarize the posterior distribution of the partition sequence, we minimize the expected posterior *variation of information* metric thus finding fully Bayesian point estimate and credibility sets, as proposed in [6].

3 Application to the Italian Causes-of-Death Networks

We here apply the dSBM model reviewed in Sect. 2 to the Italian causes-of-death data presented in Sect. 1. We set uninformative uniform prior on θ and α, corresponding to $a_\theta = b_\theta = a_\alpha = b_\alpha = 1$, and diffuse prior on η, with $a_\eta = 0.001$ and $b_\eta = 0.001$. Posterior inference is based on 40.000 MCMC samples, after a burn-in of 10.000 draws.

Figure 1 shows the adjacency matrices with rows and columns rearranged on the base of the age-specific partition point estimates obtained by minimizing the expected posterior variation of information. From this analysis, it emerges that the network nodes manifest a clustering structure smoothly evolving over age, with an increasing number of clusters that reaches its maximum at age class 50–59 and remains stable since 80–89. Younger ages are characterized by a few weakly-connected clusters, with only a few nodes manifesting a high level of connectivity. As age increases, both within-cluster and between-cluster connectivity rise, along with the overall network density. From a demographic and medical viewpoint, this pattern is reasonable, as older individuals tend to accumulate multiple chronic diseases, increasing their exposure to additional mortality risks. Across all ages, we observe a group of nodes (red cluster) lacking both internal and external connections. This weakly connected cluster is larger at younger ages but shrinks progressively with age. For instance, at birth (age 0), many neoplasms and chronic diseases fall within this cluster, as these risk factors are unlikely to affect newborns. Conversely, birth-related mortality factors, such as delivery complications or infant malformations, form a highly connected cluster at age 1 (green cluster). In subsequent networks, these causes shift to the unconnected cluster, while many neoplasms transition from the red cluster to more connected ones.

Acknowledgments. This research is funded by the MUR-PRIN 2022 project "CARONTE" (Prot. 2022KBTEBN), funded by the European Union - Next Generation EU, Mission 4, CUP: J53D23009400001.

References

1. Jemal, A., Ward, E., Hao, Y., Thun, M.: Trends in the leading causes of death in the United States, 1970–2002. JAMA **294**(10), 1255–1259 (2005)
2. Legramanti, S., Rigon, T., Durante, D., Dunson, D.B.: Extended stochastic block models with application to criminal networks. Ann. Appl. Stat. **16**(4), 2369 (2022)
3. Page, G.L., Quintana, F.A., Dahl, D.B.: Dependent modeling of temporal sequences of random partitions. J. Comput. Graph. Stat. **31**(2), 614–627 (2022)
4. Romanò, G., Castiglione, C., Durante, D.: Dependent stochastic block models for sequences of directed networks with application to causes of death co-occurences. In: Working Paper (2025)
5. Trias-Llimós, S., Permanyer, I.: Cause-of-death diversity from a multiple-cause perspective in the United States. Demography **60**(1), 73–98 (2023)
6. Wade, S., Ghahramani, Z.: Bayesian cluster analysis: point estimation and credible balls (with discussion). Bayesian Anal. **13**(2), 559–626 (2018)
7. World Health Organization (WHO): International Classification of Diseases Tenth Revision (ICD-10), Sixth Edition (2019)