Mixtures of Heterogeneous Poisson Processes for the Assessment of e-Social Activity in Mental Health

Pablo Bonilla-Escribano, David Ramírez and Antonio Artés-Rodríguez
Dept. of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain
Gregorio Marañón Health Research Institute, Spain
pbonilla@ing.uc3m.es, david.ramirez@uc3m.es, antonio@tsc.uc3m.es

Abstract

This work introduces a novel method to assess the social activity maintained by psychiatric patients using information and communication technologies. In particular, we jointly model using point processes the e-social activity patterns from two heterogeneous sources: the usage of phone calls and social and communication apps. We propose a nonhomogeneous Poisson mixture model with periodic (circadian) intensity function using a truncated Fourier series expansion, which is inferred using a trust-region algorithm, and it is able to cope with the different daily patterns of a person. The analysis of the usage of phone calls and social and communication apps of a cohort of 164 patients reveals that 25 patterns suffice to characterize their daily behavior.

1 Introduction

Mental disorders, which affect one out of four people in the world, are among the leading causes of disability [1, 2]. They represent one of the most expensive disorders to treat, and, for instance, the estimated cost of depression treatments entails more than €118 billions per year only in Europe [3]. One preponderant reason for this high treatment cost is the lack of self-awareness of the disease [4]. Indeed, if the patients' health condition worsens, yet they do not try to look for help, they may eventually suffer a relapse and may need to be hospitalized, which involves high economical and human burdens [5]. In an attempt to alleviate these burdens, a research line has recently developed tools to first gather the interactions of the patients with electronic devices, in order to establish patterns of behavior (what is referred to as digital phenotype [6]) without requiring any particular action from them. Then, data processing techniques are used to transform variations in such patterns into reliable information to diagnose diseases, evaluate their progression, or even pre-empt worsening. For instance, in [7], we studied the feasibility of detecting relapses using location data, [8] analyzed emotions using data coming from Twitter, and the work in [9] inferred online communities with interest in depression, exploiting data from social networks.

Needless to say, one of the most important interactions with devices is the one that allows to share information with other people, e.g., communication through a social network or a phone call, which we refer to as *e-social* activity. Undoubtedly, e-social activity is strongly related to actual social interactions [10], and it is especially interesting in millennials [11]. There are studies that endeavor to assess the e-social activity of psychiatric patients, such as [12], but the features that are extracted to describe it are rather daunting, since they are basically counts or proportions of the different call types. Hence, to extract meaningful conclusions out of the aforementioned counts or call proportions, which are currently taken as the indicators of the e-social activity of the patients, different machine learning techniques have been applied. For instance, in [13] decision and regression trees are used detect depression. The research of both [14] and [15] aims to predict stress, but the former uses k-nearest neighbors, SVM and PCA, while the latter employs ensembles of tree classifiers, SVM and neural networks.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

In this work, we propose a novel approach to assess the e-social activity of psychiatric patients *jointly* considering the usage of phone calls and social and communication apps. To model these types of data, it is important to consider their particular features. Concretely, they produce a set of point observations, which are composed of a start time (a timestamp), an end time (or a duration), and possibly some other features like incoming/outgoing/missed call tags, etc. Hence, the right mathematical tool to model such observations are point processes [16]. From a clinical point of view, point processes have the advantage compared to the current machine learning techniques that they directly describe the e-social activity in a way that the psychiatrist (or even the patient) can understand it without requiring any kind of black-box algorithm.

In order to account for different factors that may influence the patients' routines, such as holidays, weekdays vs. weekends, etc., mixture models may be employed to parametrize the (conditional) intensity function. Although mixtures of Poisson point processes have been successfully applied in other fields, like vehicular accident data [17], insurance claims [18], economy [19, 20], RNA sequencing [21], or earthquake modeling [22], to the best of our knowledge, our work in [23] is the only study where mixtures of Poisson processes are used for the assessment of e-social activity of psychiatric patients. Nonetheless, the interpretability of the model therein is hindered by the fact that each source (i.e. the usage of phone calls and social and communication apps) was analyzed independently; that is, using two hidden variables that govern different intensity functions for the two sources. Indeed, the presence of different types of intensities to model the sources may be misleading. To alleviate it, we propose a model that extracts patterns taking into account heterogeneous sources at inference time. Further, since there is strong evidence supporting that human activities, such as socializing, follow a circadian rhythm, [24, 25], we propose to parametrize each of the components of the mixture model by means of periodic functions using a truncated Fourier series expansion with order to $K = 3$, which is consistent with the existing literature [26].

2 Model

In this section, we introduce a novel point process whose intensity is based on a mixture of intensities describing heterogeneous sources and, moreover, only one intensity may be active on each day for each source, in such a way that each intensity can be associated with a particular *daily* pattern. That is, we consider a model wherein there are u' different sources, and for each of them, there is a known number of (conditional) intensities, N, generating them. In addition, those (conditional) intensities remain active during a time slot of duration $T = 24h$, and could only change at times $T \times m$, where $m = 1, 2, \ldots, m'$. Here, m' is the number of observation days. The aforementioned model can be defined by means of a single hidden variable $z(t)$ whose value indicates which *set* of (conditional) intensities has generated all the observed events at time t from all sources. Moreover, defining $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_N]^T$ as the vector of prior probabilities on $z(t)$, the likelihood function, $L_{\text{heterogeneous}}$, is given by

$$L_{\text{heterogeneous}} = \prod_{m=1}^{m'} \left\{ \prod_{l=1}^{N} \left[\pi_l \prod_{u=1}^{u'} \left(\left(\prod_{n \in \mathcal{N}_u(m)} f^*_{l,u}(t_{m,n}) \right) \times \left(1 - F^*_{l,u}(mT) \right) \right) \right]^{\mathbb{I}\left(z(t_m)=l \right)} \right\}, \tag{1}$$

where $f^*_{l,u}(t)$ is the (conditional) density that an event happens at time t when the lth mixture component is active for the uth source, and $F^*_{l,u}(t)$ is the corresponding cumulative distribution function (CDF). In addition, $t_{m,n}$ is the time instant of the nth observation within the mth day, and $\mathcal{N}_u(m)$ is the set of indexes of the observations contained in the mth day for the uth source. Finally, a single subscript represents any time instant within the mth day, i.e., t_m.

To proceed, (1) is rewritten in terms of (conditional) intensities, following [27]. Then, the new expression is optimized using the EM algorithm [28, 29], which has a straightforward E-step that mainly involves the computation of the responsibilities $r_{m,l}$. On the other hand, in the M-step, the optimization with respect to each (conditional) intensity of each source is independent, and taking into account the parametric form presented in [23], the ML estimator of each $\lambda^*_{l,u}(t) = \mathbf{d}^T_{l,u}\mathbf{T}(t)\mathbf{d}_{l,u}$

is the solution of the minimization problem [27]

$$\underset{\mathbf{d}_{l,u}}{\operatorname{argmin}} \underbrace{\mathbf{d}_{l,u}^T \left(\sum_{m=1}^{m'} r_{m,l} \overline{\mathbf{T}}_m \right) \mathbf{d}_{l,u} - \left[\sum_{m=1}^{m'} \sum_{n \in \mathcal{N}_u(m)} r_{m,l} \log \left(\mathbf{d}_{l,u}^T \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u} \right) \right]}_{J(\mathbf{d}_{l,u})}, \tag{2}$$

where $\mathbf{d}_{l,u}$ is the vector of Fourier coefficients,

$$\mathbf{T}(t) = \begin{bmatrix} \mathbf{C}(t) & \mathbf{S}(t) \\ \mathbf{S}^T(t) & \mathbf{C}(t) \end{bmatrix}, \qquad \overline{\mathbf{T}}_m = \begin{bmatrix} \overline{\mathbf{C}}_m & \overline{\mathbf{S}}_m \\ \overline{\mathbf{S}}_m^T & \overline{\mathbf{C}}_m \end{bmatrix},$$

with the blocks of the matrices being

$$\overline{\mathbf{C}}_m = \int_{(m-1)T}^{mT} \mathbf{C}(t)\, dt, \qquad \overline{\mathbf{S}}_m = \int_{(m-1)T}^{mT} \mathbf{S}(t)\, dt,$$

and, defining $\mathbf{U}_k$ as a Toeplitz matrix whose entries on the kth diagonal are 1 and 0 elsewhere,

$$\mathbf{C}(t) = \mathbf{I}_{K+1} + \sum_{k=1}^{K} \cos\left(\frac{2\pi k}{T} t \right) \left(\mathbf{U}_k + \mathbf{U}_k^T \right), \qquad \mathbf{S}(t) = \sum_{k=1}^{K} \sin\left(\frac{2\pi k}{T} t \right) \left(\mathbf{U}_k - \mathbf{U}_k^T \right).$$

It is easy to see that the optimization problem in (2) is not convex [30], thus requiring the use of non-convex techniques. In particular, we propose to use the trust-region method in [31], which is based on the interior-reflective Newton's method. This technique requires both the gradient and the Hessian of the cost function, which are computed next. Using [32], we obtain the required gradient of $J(\mathbf{d}_{l.u})$, which is given by

$$\nabla_{\mathbf{d}_{l,u}} J(\mathbf{d}_{l,u}) = -2 \left\{ \sum_{m=1}^{m'} \sum_{n \in \mathcal{N}_u(m)} r_{m,l} \frac{1}{\mathbf{d}_{l,u}^T \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u}} \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u} \right\} + 2 \left(\sum_{m=1}^{m'} r_{m,l} \overline{\mathbf{T}}_m \right) \mathbf{d}_{l,u}. \tag{3}$$

Similarly, the Hessian of the cost function is

$$\begin{aligned} \mathbf{H} J(\mathbf{d}_{l,u}) = \Bigg\{ \sum_{m=1}^{m'} \sum_{n \in \mathcal{N}_u(m)} r_{m,l} \Bigg[& \left(\frac{2}{\mathbf{d}_{l,u}^T \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u}} \right)^2 \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u} \mathbf{d}_{l,u}^T \mathbf{T}(t_{m,n}) \\ & - \frac{2}{\mathbf{d}_{l,u}^T \mathbf{T}(t_{m,n}) \mathbf{d}_{l,u}} \mathbf{T}(t_{m,n}) \Bigg] \Bigg\} + 2 \left(\sum_{m=1}^{m'} r_{m,l} \overline{\mathbf{T}}_m \right). \end{aligned} \tag{4}$$

3 Results

Fig. 1 shows the 25 intensity pairs that are relevant enough (according to the AIC criterion [33]) in order to explain the e-social activity of the 164 patients, which involves 385 213 entries in the usage log of social and communication apps (WhatsApp, Twitter...) and 280 239 phone calls. Actually, as shown in Fig. 2, none of the patients activated all the patterns, but just a limited amount of them. This shows that their e-social activity can be understood in a succinct way thanks to this method. With regard to the patterns in Fig. 1, they are ordered in increasing number of phone calls, and it is interesting to notice that (except for pairs 16, 17, 18, 24 and 25, where phone calls are the preferred method; and 19 and 23, where the usage is similar) the activity on social and communication apps supersedes and is more complex than that of the phone calls.

The pair 10 is interesting since it reveals a nocturnal pattern in the usage of social and communication apps, but low activity making phone calls. It is not surprising that the method only finds nocturnal patterns in the usage of social and communication apps and not in phone calls. Indeed, the patients can use social networks or send instant messages at nocturnal times without disturbing their contacts.

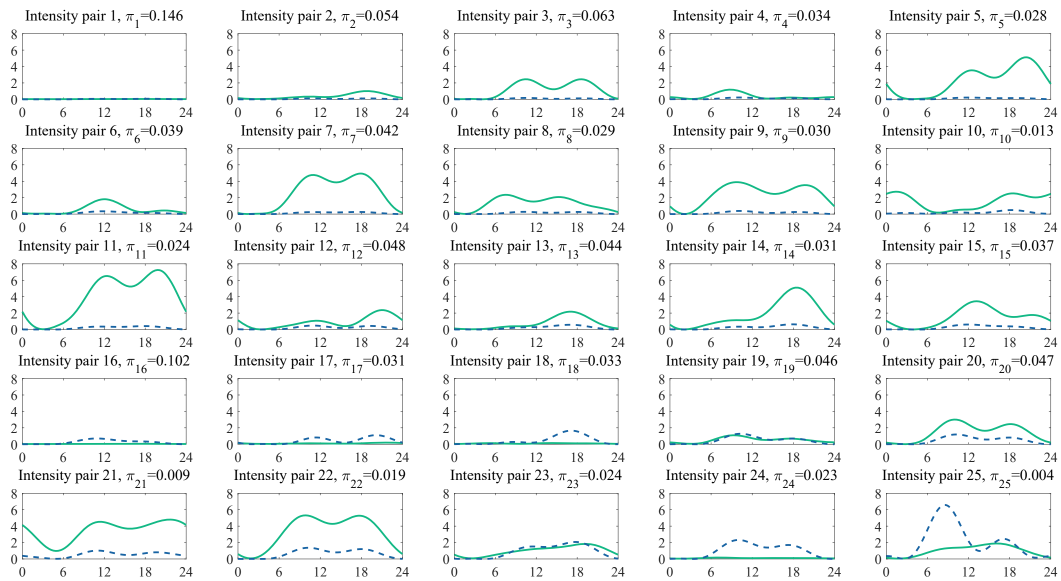

Figure 1: Intensity pairs obtained by the proposed method from the usage of calls (dashed blue) and social and communication apps (solid green) by 164 psychiatric patients. The abscissa represents the hour of the day, while the ordinate represents the intensity in events/h.

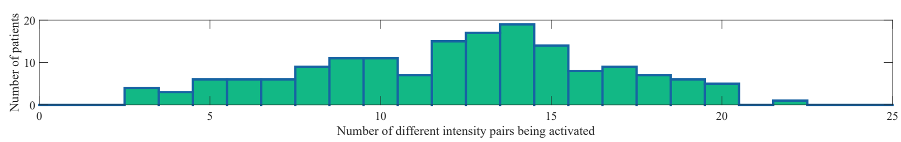

Figure 2: Histogram of the number of different intensities pairs activated by the patients

However, they would require to know other people who do not regularly sleep during at night in order to make phone calls with them. Moreover, many intensities show a two-period activity, where the social activities are concentrated around morning and evening (like the pairs 20 and 22). There are also patterns of unimodal activity, such as the phone usage of the intensity pair 18. The intensity pair 24 requires more attention, since it shows the e-social activity of those patients uninterested in social networks and related platforms, who make a continuous use of phone calls in order to keep in contact with their peers.

With the exception of the pairs 19 through 23, the shape of the intensity pairs are not similar, since one of the communication methods tends to dominate over the other. For instance, the pair 25 shows that there are patients that use social and communication apps in a moderate, uniform way at diurnal times, but they make extensive use of mobile phones at 10 a.m., and moderate use at 6 p.m. This shows that e-social is a complex, heterogeneous phenomenon, which cannot be properly understood assuming that the overall activity factorizes over all the sources.

4 Conclusions

We have presented a novel tool that can be applied to characterize the e-social activity of psychiatric patients, succinctly integrating information coming from heterogeneous sources. This tool is based on a Poisson process mixture model, which requires a very limited number of intensity sets in order to explain the behavior of large amounts of e-social data from psychiatric patients.

From the clinical point of view, we have found a set of interpretable patterns analyzing data coming from the usage of phone calls and social and communication apps. For instance, the method has discovered patterns of nocturnal activity and poor social interactions. Moreover, since social activity is an integral part of the patient's behavior, we believe that the methods presented in this work could serve as a basis to help with the diagnosis of mental illnesses and to design the treatment.

Acknowledgments

This work was supported by the Ministerio de Ciencia, Innovación y Universidades under grant TEC2017-92552-EXP (aMBITION), by the Ministerio de Ciencia, Innovación y Universidades, jointly with the European Commission (ERDF), under grants TEC2017-86921-C2-2-R (CAIMAN) and RTI2018-099655-BI00 (CLARA), and by The Comunidad de Madrid under grant Y2018/TCS-4705 (PRACTICO-CM).

References

[1] World Health Organization, "The world health report 2001 - Mental health: New understanding, new hope," 2001.

[2] H. A. Whiteford, A. J. Ferrari, L. Degenhardt, V. Feigin, and T. Vos, "The global burden of mental, neurological and substance use disorders: An analysis from the global burden of disease study 2010," *PloS one*, vol. 10, no. 2, 2015.

[3] P. Sobocki, B. Jönsson, J. Angst, and C. Rehnberg, "Cost of depression in Europe." *Journal of Mental Health Policy and Economics*, vol. 9, no. 2, pp. 87–98, 2006.

[4] X. F. Amador, M. Flaum, N. C. Andreasen, D. H. Strauss, S. A. Yale, S. C. Clark, and J. M. Gorman, "Awareness of illness in schizophrenia and schizoaffective and mood disorders," *Archives of General Psychiatry*, vol. 51, no. 10, pp. 826–836, 1994.

[5] B. L. Svarstad, T. I. Shireman, and J. Sweeney, "Using drug claims data to assess the relationship of medication adherence with hospitalization and costs," *Psychiatric Services*, vol. 52, no. 6, pp. 805–811, 2001.

[6] S. H. Jain, B. W. Powers, J. B. Hawkins, and J. S. Brownstein, "The digital phenotype," *Nature Biotechnology*, vol. 33, no. 5, p. 462, 2015.

[7] S. Berrouiguet, D. Ramírez, M. L. Barrigón, P. Moreno-Muñoz, R. Carmona, E. Baca-García, and A. Artés-Rodríguez, "Combining continuous smartphone native sensors data capture and unsupervised data mining techniques for behavioral changes detection: The feasibility study of the Evidence Based Behavior (eB2) platform," *JMIR mHealth and uHealth*, vol. 6, no. 12, 2018.

[8] M. E. Larsen, T. W. Boonstra, P. J. Batterham, B. O'Dea, C. Paris, and H. Christensen, "We feel: Mapping emotion on Twitter," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1246–1252, July 2015.

[9] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh, "A framework for classifying online mental health-related communities with an interest in depression," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, pp. 1008–1015, July 2016.

[10] N. B. Ellison, C. Steinfield, and C. Lampe, "The benefits of Facebook "friends:" Social capital and college students' use of online social network sites," *Journal of Computer-Mediated Communication*, vol. 12, no. 4, pp. 1143–1168, 2007.

[11] W. Youyou, M. Kosinski, and D. Stillwell, "Computer-based personality judgments are more accurate than those made by humans," *Proceedings of the National Academy of Sciences*, vol. 112, no. 4, pp. 1036–1040, 2015.

[12] D. Ben-Zeev, R. Brian, R. Wang, W. Wang, A. T. Campbell, M. S. Aung, M. Merrill, V. W. Tseng, T. Choudhury, M. Hauser *et al.*, "Crosscheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse." *Psychiatric Rehabilitation Journal*, vol. 40, no. 3, p. 266, 2017.

[13] M. N. Burns, M. Begale, J. Duffecy, D. Gergle, C. J. Karr, E. Giangrande, and D. C. Mohr, "Harnessing context sensing to develop a mobile intervention for depression," *Journal of Medical Internet Research*, vol. 13, no. 3, 2011.

[14] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pp. 671–676.

[15] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland, "Daily stress recognition from mobile phone data, weather conditions and individual traits," *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 477–486, 2014.

[16] J. F. C. Kingman, *Poisson processes*. Clarendon Press, 1992, vol. 3.

[17] A. T. Ihler and P. Smyth, "Learning time-intensity profiles of human activity using non-parametric Bayesian models," *Advances in Neural Information Processing Systems*, pp. 625–632, 2007.

[18] G. O. Brown and W. S. Buckley, “Experience rating with Poisson mixtures,” *Annals of Actuarial Science*, vol. 9, no. 2, pp. 304–321, 2015.

[19] S. Faria and F. Gonçalves, “Financial data modeling by Poisson mixture regression,” *Journal of Applied Statistics*, vol. 40, no. 10, pp. 2150–2162, 2013.

[20] M. Burda, M. Harding, and J. Hausman, “A Poisson mixture model of discrete choice,” *Journal of Econometrics*, vol. 166, no. 2, pp. 184–203, 2012.

[21] H. Wu, Z. Qin, and Y. Zhu, “PM-Seq: Using finite Poisson mixture models for RNA-Seq data analysis and transcript expression level quantification,” *Statistics in Biosciences*, vol. 5, no. 1, pp. 71–87, 2013.

[22] A. C. Micheas and J. Chen, “sppmix: Poisson point process modeling using normal mixture models,” *Computational Statistics*, vol. 33, no. 4, 2018.

[23] P. Bonilla-Escribano, D. Ramírez, A. Sedano-Capdevila, J.-J. Campana-Montes, E. Baca-García, P. Courtet, and A. Artés-Rodríguez, “Assessment of e-social activity in psychiatric patients,” *IEEE journal of biomedical and health informatics*, 2019.

[24] W. Huang, K. M. Ramsey, B. Marcheva, and J. Bass, “Circadian rhythms, sleep, and metabolism,” *The Journal of Clinical Investigation*, vol. 121, no. 6, pp. 2133–2141, 2011.

[25] A. Sorokin, A. Maksimov, and J. Jermain, “Human circadian rhythm synchronization by social timers: The role of motivation: IV. Individual features of the free-running 24-hour sleep-wake cycle under simulated conditions of vital activity,” *Human Physiology*, vol. 26, no. 1, pp. 41–47, 2000.

[26] T. Aledavood, E. López, S. G. Roberts, F. Reed-Tsochas, E. Moro, R. I. Dunbar, and J. Saramäki, “Daily rhythms in mobile telephone communication,” *PloS one*, vol. 10, no. 9, 2015.

[27] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: Volume II: General theory and structure*. Springer-Verlag New York, 2008.

[28] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[29] C. F. J. Wu, “On the convergence properties of the EM algorithm,” vol. 11, no. 1, pp. 95–103, 1983.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[31] T. F. Coleman and Y. Li, “An interior trust region approach for nonlinear minimization subject to bounds,” *SIAM Journal on Optimization*, vol. 6, no. 2, pp. 418–445, 1996.

[32] J. R. Magnus and H. Neudecker, “Matrix differential calculus with applications to simple, Hadamard, and Kronecker products,” *Journal of Mathematical Psychology*, vol. 29, no. 4, pp. 474–492, 1985.

[33] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike information criterion statistics*, ser. Mathematics and its Applications. Springer Netherlands, 1986.