http://doi.org/10.31234/osf.io/TBD

Modeling latent topics in social media using Dynamic Exploratory Graph Analysis: The case of the right-wing and left-wing trolls in the 2016 US elections

Hudson Golino[1*], Alexander P. Christensen[2], Robert Moulder[1], Seohyun Kim[1], and Steven M. Boker[1]

ABSTRACT

The past few years were marked by increased online offensive strategies perpetrated by state and non-state actors to promote their political agenda, sow discord and question the legitimacy of democratic institutions in the US and Western Europe. In 2016 the US congress identified a list of Russian state-sponsored Twitter accounts that were used to try to divide voters on a wide range of issues. Previous research used Latent Dirichlet Allocation (LDA) to estimate latent topics in data extracted from these accounts. Howerver, LDA is has characteristics that may pose significant limitations to be used in data from social media: the number of latent topics must be specified by the user, interpretability can be difficult to achieve, and it doesn't model short-term temporal dynamics. In the current paper we propose a new method to estimate latent topics in texts from social media termed Dynamic Exploratory Graph Analysis (DynEGA). We compare DynEGA and LDA in a Monte-Carlo simulation in terms of their capacity to estimate the number of simulated latent topics. Finally, we apply the DynEGA method to a large dataset with Twitter posts from state-sponsored right- and left-wing trolls during the 2016 US presidential election. The results show that DynEGA is substantially more accurate to estimate the number of simulated topics than several different LDA algorithms. Our empirical example shows that DynEGA revealed topics that were pertinent to several consequential events in the election cycle, demonstrating the coordinated effort of trolls capitalizing on current events in the U.S. This demonstrates the potential power of our approach for revealing temporally relevant information from qualitative text data.

***Keywords*:** text mining, latent topic analysis, network models, dynamics, time embedding

Compiled: April 3rd 2020

Note: This article has been submitted to *Psychometrika* for their special issue on "Network Psychometrics in Action – Applications and Case Studies."

Introduction

The past few years were marked by increased online offensive strategies perpetrated by state and non-state actors to promote discord and question the legitimacy of democratic institutions in the US and Western Europe (Taddeo, 2017; Ziegler, 2018). These offensive strategies ranged from traditional cyber-attacks (e.g. denial-of-service, data leaking, and application compromising; Hernandez-Suarez et al., 2018) to information warfare—a set of tactics and operations involving the protection, manipulation, degradation, and denial of information (Libicki, 1995). The goals of information warfare are to attack adversary knowledge or beliefs (Szafranski, 1995), fabricate false or distorted stories, generate opposition movements, and destabilize adversaries (Ziegler, 2018). In the past decade, the increase in the number of people using social media platforms (such as Twitter) and sharing content online (over 1 billion posts per month around the world; Hernandez-Suarez et al., 2018) has led to increased gains from information operations on both scale and impact.

A recent notable information operation was the covert online activities in social media to influence the public opinion and voters in the U.S. during the 2016 campaign, with attacks occurring before and during the electoral process (Linvill & Warren, 2018). Social media accounts linked to the Internet Research Agency (IRA), based in Russia, were used to sow discord in the U.S. political system, using trolls and robots that masqueraded as American citizens to try to divide voters on a wide range of issues (Linvill & Warren, 2018). Qualitative analysis of the content published by IRA-linked Twitter accounts have been conducted elsewhere (see: Linvill, Boatwright, Grant, & Warren, 2019), providing relatively little insight despite the large amount of information posted by the trolls (almost 3 million tweets from 2,848 Twitter handles).

Quantitative analysis of texts, however, can be very useful in understanding the strategies used in online intelligence operations. Llewellyn, Cram, Favero, and Hill (2018) used topic modeling to understand the content disseminated by IRA trolls in Twitter related to the UK-EU Referendum debate (*Brexit*), and found that the topics revolved around the economy, European Union affairs, voting and external politics, and various other topics. Ghanem, Buscaldi, and Rosso (2019) used a combination of latent topic modeling and stylistic text analysis (i.e., the use of textual features to differentiate authorship) to predict IRA-linked Twitter

[1] University of Virginia, Charlottesville, VA, USA
[2] University of North Carolina at Greensboro, Greensboro, NC, USA
[*] Corresponding author: hfg9s@virginia.edu

How to cite: Golino, H., Christensen, A. P., Moulder, R., Kim, S., & Boker, S. M. (2020). Modeling latent topics in social media using Dynamic Exploratory Graph Analysis: The case of the right-wing and left-wing trolls in the 2016 US elections. *PsyArXiv*.

accounts. The results of Ghanem et al. (2019) show that modeling latent topics combined with surface text features (stylistic analysis) can improve the performance of predictive models.

In another study focused on IRA-linked accounts, Zannettou, Caulfield, De Cristofaro, et al. (2019) compared IRA-linked troll tweets with a random set of Twitter users and investigated their ability to spread and make news content viral. The authors discovered that while the IRA-linked trolls are active for long periods of time and reach a substantial number of Twitter users, their effect on making content go viral was limited, with the exception of news published by a Russian state-sponsored news outlet (Zannettou, Caulfield, De Cristofaro, et al., 2019). In a separate study, Zannettou, Caulfield, Setzer, et al. (2019) compared Twitter and Reddit users identified as Russian and Iranian trolls and found, among other things, that Russian trolls were spreading pro-Trump content while Iranian trolls were spreading anti-Trump content.

In all of the papers mentioned above (Ghanem et al., 2019; Llewellyn et al., 2018; Zannettou, Caulfield, De Cristofaro, et al., 2019; Zannettou, Caulfield, Setzer, et al., 2019), *Latent Dirichlet Allocation* (LDA; Blei, Ng, & Jordan, 2003) was used to estimate latent topics from posts on social media platforms. LDA is one of the most well-known and widely used topic modeling techniques, yet it has several limitations: (1) the number of latent estimated topics must be specified by the user, (2) interpretability can be difficult to achieve (e.g., probabilities for each topic must be interpreted), (3) assumes the topics are uncorrelated, and (4) doesn't model short-term temporal dynamics. The estimation of the number of latent topics is a problem akin to the dimensionality assessment problem in psychometrics. Although some algorithms were developed to check the optimal number of latent topics estimated via LDA (Arun, Suresh, Veni Madhavan, & Narasimha Murthy, 2010; Cao, Xia, Li, Zhang, & Tang, 2009; Deveaud, SanJuan, & Bellot, 2014), their implementation involves multiple steps and was not widely available in software until very recently (Nikita, 2019), which may help explain why researchers choose the number of topics in LDA arbitrarily.

In terms of interpretability, topic modeling methods are statistical tools in which numerical distributions must be explored to generate a meaningful interpretation of the results (Chaney & Blei, 2012). In LDA and most topic modeling techniques, the output of the models do not provide enough information to generate a straight-forward interpretation of the latent topics (Chaney & Blei, 2012), and the researchers must check the distribution of word probabilities per topic in order to make sense of them. In real-world data, interpreting the distribution of word probabilities per latent topic is not just cumbersome, but sometimes the words with the highest probabilities in each topic are very similar, making the interpretation of the content of the topics in LDA (and other topic modeling methods) a challenge.

Some of the limitations of LDA, such as assuming the topics are uncorrelated and not modeling temporal dynamics, were addressed previously. For example, Blei and Lafferty (2007) developed a topic modeling technique that allows the topics to be correlated, Blei and Lafferty (2006) proposed a new topic modeling technique to model temporal dynamics, and Glynn, Tokdar, Howard, and Banks (2019) proposed a Bayesian extension of the dynamic topic model (Glynn et al., 2019). Despite the usefulness of these correlated and dynamic topic models, they share the above mentioned limitations of LDA: the need to specify the number of topics *a priori* and the difficulty to interpret the results. Furthermore, the computation time can be very expensive and the availability of R (R Core Team, 2018) packages implementing these techniques are limited (if existent at all). The Bayesian dynamic topic model proposed by Glynn et al. (2019), for example, takes several hours to run *a single topic model estimation* on a personal computer and although the correlated topic model of Blei and Lafferty (2007) is implemented in R (see Hornik & Grun, 2011), the dynamic topic model of Blei and Lafferty (2006) is not.

In the current paper we propose a new (and fast) method to estimate latent dimensions (e.g., factors, topics) in multivariate time series termed *Dynamic Exploratory Graph Analysis* (DynEGA). The DynEGA technique can be used to estimate the latent structure of topics published in social media (using time series of word frequencies), improving our capacity to understand the strategy used by accounts created as tools of information warfare. The DynEGA approach uses time delay embedding to pre-process each variable (e.g., time series of words counts), and estimates the derivatives from each variable using *Generalized Local Linear Approximation* (GLLA; Boker, Deboek, Edler, & Keel, 2010). Finally, a network psychometrics approach for dimensionality assessment termed *Exploratory Graph Analysis* (EGA; Golino & Epskamp, 2017; Golino et al., 2020) is used to identify clusters of variables that are changing together (i.e., dynamic latent factors or dynamic latent topics in the case of text data). The DynEGA approach automatically estimates the number of latent factors or topics and their short-term temporal dynamics with the results displayed as a network plot, which facilitates their interpretation. Importantly, the DynEGA approach does not assume the factors or topics are uncorrelated (e.g., LDA and other commonly applied topic modeling techniques), and it can accommodate different time scales and estimate the latent structure at different levels of analysis (i.e., population, groups, and individuals).

After describing the DynEGA method, we compare it to LDA in a brief simulation study where the latent topics are simulated using the direct autoregressive factor score model (**DAFS**; Engle & Watson, 1981; Nesselroade, McArdle, Aggen, & Meyers, 2002), which is characterized by the autoregressive structure of the latent dimensions (Nesselroade et al., 2002). In the simulation, LDA is implemented using different estimation methods via the `topicmodels` package (Hornik & Grun, 2011) and the number of latent topics is verified using AIC, BIC and the algorithms developed by Arun et al. (2010), Cao et al. (2009) and Deveaud et al. (2014).

Finally, we apply the DynEGA method to the Twitter data published by Linvill and Warren (2018), which contains posts from IRA-linked accounts that were identified as right- and left-wing trolls. The goals of the current paper are to introduce the new dynamic EGA model, verify its suitability to estimate latent topics in a brief simulation study, and investigate the strategies used by right- and left-wing trolls to sow discord in the U.S. political system. Being able to identify the communication strategies used by the IRA can potentially enhance our capacity to understand online intelligence operations, which are likely part of information warfare efforts perpetrated by both state and non-state actors. We've implemented the DynEGA method into the EGAnet package for the R software environment (Golino & Christensen, 2019; Golino et al., 2020). All code used in the current paper are available in an online repository at the Open Science Framework platform for reproducibility purposes (see: https://osf.io/4ya6x/?view_only=b6078b404e3049818b359ae0d514f966).

Exploratory Graph Analysis: a (very) brief overview

The origins of network models in psychology can be traced back to the seminal work of Cattell in the mid-60's (Boker, 2018; Cattell, 1965) and less explicitly to the proposition of image structural analysis by Guttman (1953). It gained more traction, however, after the publication of the mutualism model of intelligence (Van Der Maas et al., 2006) and the proposition of the network perspective of psychopathological constructs (Borsboom, 2008; Borsboom & Cramer, 2013; Cramer, Waldorp, Van Der Maas, & Borsboom, 2010; Fried et al., 2017) as well as being employed in clinical (Bork, Borkulo, Waldorp, Cramer, & Borsboom, 2018), cognitive (Golino & Demetriou, 2017; Van Der Maas, Kan, Marsman, & Stevenson, 2017), social (Dalege, Borsboom, Harreveld, Waldorp, & Maas, 2017), and many other areas of psychology (Epskamp, Rhemtulla, & Borsboom, 2017).

The rapid developments of network modeling in psychology spawned a new subfield of quantitative psychology termed *network psychometrics* (Epskamp, 2018). In these models, nodes (e.g., circles) represent variables and edges (e.g., lines) represent associations between the nodes. Under this framework, Golino and Epskamp (2017) proposed the use of network psychometrics as a method for dimensionality assessment and termed this novel approach EGA. Unlike other methods, EGA produces a visual guide—network plot—that not only indicates the number of dimensions to retain, but also which nodes (e.g., items) cluster together and their level of association. Simulation studies have shown that EGA presents comparable or better accuracy than the state-of-the-art parallel analysis technique when estimating the number of simulated factors (Christensen, 2020; Golino & Epskamp, 2017; Golino et al., 2020).

The EGA approach currently uses two network estimation methods (for a review, see Golino et al., 2020): graphical *lasso* (Friedman, Hastie, & Tibshirani, 2008) and triangulated maximally filtered graph (TMFG; Massara, Di Matteo, & Aste, 2016). After the network is estimated, an algorithm for community (cluster) detection in weighted networks is used (Walktrap; Pons & Latapy, 2006). The next sections will briefly introduce the network estimation methods and the community detection algorithm used in EGA.

Graphical lasso

The graphical *lasso* (*glasso*; Friedman et al., 2008) is the most commonly applied network estimation method in the psychometric network literature. Networks estimated using the *glasso* are a Gaussian Graphical Model (GGM; Lauritzen, 1996), where edges represent partial correlations between variables after conditioning on all other variables in the network. The least absolute shrinkage and selection operator (*lasso*; Tibshirani, 1996) is used to control for spurious relationships and shrink coefficients to zero, generating a sparse network and preventing over-fitting of the data. The *glasso* procedure can be controlled in a way to generate multiple networks, with different levels of regularization (i.e., from a fully connected network to a fully unconnected network). This approach is termed the *glasso path*, in which *glasso* is run for *n* values of the tuning parameter λ. For each estimated network, the extended Bayesian information criterion (EBIC; Chen & Chen, 2008) is computed and the graph with the best EBIC is selected (Epskamp et al., 2018; Epskamp & Fried, 2018; Foygel & Drton, 2010). The EBIC has a hyperparameter (γ) that controls the severity of the model selection (i.e. that controls how much the EBIC prefers simpler models; Epskamp & Fried, 2018). Most commonly, γ is set to 0.5 (Foygel & Drton, 2010), although greater sensitivity (true positive proportion) can be gained from lower γ values (e.g., $\gamma = 0$) at the cost of specificity (true negative proportion; Williams & Rast, 2019). In EGA, the network estimation starts with $\gamma = 0.5$, but if the resulting network has disconnected nodes, then γ is set to 0.25, and then to 0 if the issue persists. When γ is zero, EBIC equals the Bayesian information criterion (Foygel & Drton, 2010).

Triangulated Maximally Filtered Graph

The triangulated maximally filtered graph (TMFG; Massara et al., 2016) is another network construction method that has been used in the psychometric network literature (e.g., Christensen, Kenett, Aste, Silvia, & Kwapil, 2018; Golino et al., 2020). The TMFG algorithm constructs the network using a structural constraint on the number of zero-order correlations that can be included in the network (i.e., $3n-6$; where n is the number of variables). Construction begins by identifying the four variables with largest sum of correlations to all other variables and connects them to each other. Then, variables are iteratively added to the network based on the largest sum of three correlations to nodes already in the network. The result is a fully connected network of 3- and 4-node cliques (i.e., sets of connected nodes). This 3- and 4-node clique structure can be directly associated with the inverse covariance matrix, resulting in a GGM (Barfuss, Massara, Di Matteo, & Aste, 2016).

Walktrap Algorithm

EGA uses the Walktrap community detection algorithm (Pons & Latapy, 2006) to determine the number of *communities* (factors or topics) in the network (Golino & Epskamp, 2017). The Walktrap algorithm uses a process known as *random walks* (or a stochastic number of edges from a certain node), which tend to get "trapped" in densely connected parts of the network. The number of steps can be specified by the user; however, EGA uses the default of 4, which has been shown to be optimal for the number of variables typically used in psychological research (Christensen, 2020; Gates, Henry, Steinley, & Fair, 2016). In the random walk process, the likelihood of a step to another node is determined by the structural similarity between the nodes and communities, which defines a distance.

These distances are used in an agglomerative hierarchical clustering algorithm approach, which is then subjected to merging (i.e., merging two clusters that minimize the mean of the squared distances). During the merging process, the adjacent clusters' distances are updated to reflect the new distances between the clusters. Throughout this process, a metric to assess the quality of the partitions is used to help capture community structures at different scales.

Dynamic Exploratory Graph Analysis

Dimensionality assessment is common in psychology, but also has a significant use in data mining, especially in the subfield of text mining. Text mining is a data-driven, exploratory method used to find patterns and trends in large data sets of texts, enabling the transformation of unorganized text into succinct knowledge (Ananiadou & McNaught, 2006). It is epistemologically compatible with content analysis, making it possible to collect, maintain, interpret, and discover relevant information hidden in texts in a systematic and efficient way (Singh, Hu, & Roehl, 2007). Recently, EGA was used in combination with text mining to estimate latent topics in texts, showing promising results (Kjellstrom & Golino, 2019).

The current implementation of EGA, however, limits its application to data collected to a single time-point (i.e., cross-sectional data). Kjellstrom and Golino (2019), for example, used text data from single interviews made with multiple adults about their conceptions of health. To enable the identification of latent structures in texts from social media, the EGA technique needs to be expanded to accommodate short-term temporal dynamics. This, in our view, would provide a more valid way to estimate topics in texts that are produced in series (such as posts in Twitter). People use words to communicate their ideas and feelings and thoughts, with groups of words indicating the underlying content of the text (i.e., the topic). If the text data comes from a single time point, the topics can be estimated using the variance-covariance matrix of the words (or the inverse of the variance-covariance matrix, as in the *glasso* version of EGA; Kjellstrom & Golino, 2019). However, in the case of texts that are written on several different occasions, the temporal dynamics should be accounted for by the topic model, otherwise it may generate significant bias in the estimation of the underlying latent topics.

In texts published on Twitter, people may use several words, for example, to talk about the topic "violence" on one occasion (time or t). In the next occasion ($t + 1$), the same person may use a different set of words to express their feelings about "gun control," and yet in a subsequent occasion may use different words to communicate their views about the "mainstream media." Across time, one person can write about several different topics, using words that may or may not be the same on each occasion. Instead of modeling the covariance of words without taking time into consideration (a cross-sectional approach), a more ecologically valid way to understand topics published on Twitter is to model how words are varying together across time, capturing the short-term dynamics of the texts.

One way to address this problem is proposed as follows: A collection of texts (corpus) from one single individual (e.g., a Twitter account) over N discrete time points can be represented as a document-term matrix (DTM) in which each unique word is a column and each observation (e.g., a post on Twitter) is a row in the document-term matrix. The DTM is, therefore, a $N \times U$ matrix, where N is the number of time points and U is the number of unique words used in the entire collection of texts (corpus). The values of the DTM cells are the frequency of the words.

Since each column of the document-term matrix represents a time series of the word frequency, $W = \{w_1, w_2, ..., w_N\}$, each time series can be transformed into a time delay embedding matrix $\mathbf{X}^{(n)}$, where n is the number of embedding dimensions. A time delay embedding matrix is used to reconstruct the attractor of a dynamical system using a single sequence of observations (Takens, 1981; Whitney, 1936). An attractor contains useful information about the dynamical system such as a series of values to which a system tends toward based on a set of starting conditions. In many empirical situations, however, the collection of possible system states (phase-space) and the equations governing the system are unknown. In such situations, attractor reconstruction techniques can be used as a means to reconstruct the phase-space dynamics using, for example, only a single time series with observable values.

In the time delay embedding matrix, each row is a phase-space vector (Rosenstein, Collins, & De Luca, 1993):

$$X = [X_1 \; X_2 \; ... \; X_M]' \tag{1}$$

Where X_i is the state of the system at discrete time i and is given by:

$$X_i = \left[x_i \; x_{i+\tau} \; ... \; x_{i+(n-1)\tau}\right] \tag{2}$$

where τ is the number of observations to offset successive embeddings (i.e., lag or reconstruction delay) and n is the embedding dimension. The time-delay embedding matrix is a $M \times n$ matrix, where $M = N - (n-1)\tau$ and N is the number of observations.

Suppose that W_1 is a column in a given document-term matrix representing the time series (of the frequency) of the word *gun*, from time $t = 1$ to $t = 10$, so that $W_1 = \{5, 6, 7, ..., 14\}$. The frequencies of the word *gun* in this example are way beyond what one finds in textual data, especially from social media platforms, but the goal of the example is to help the reader understand how time delay embedding works. Transforming the time series W_1 into a time delay embedding matrix with five embedding dimensions and $\tau = 1$ generates in the following matrix:

$$\mathbf{X}^{(5)} = \begin{bmatrix} 5 & 6 & 7 & 8 & 9 \\ 6 & 7 & 8 & 9 & 10 \\ 7 & 8 & 9 & 10 & 11 \\ 8 & 9 & 10 & 11 & 12 \\ 9 & 10 & 11 & 12 & 13 \\ 10 & 11 & 12 & 13 & 14 \end{bmatrix} \tag{3}$$

Once every time series of word frequency (columns) of the document-term matrix is transformed into a time-delay embedding matrix $\mathbf{X}^{(n)}$, derivatives can be estimated using GLLA (Boker et al., 2010; Deboeck, Montpetit, Bergeman, & Boker, 2009).

GLLA is a technique that can be used to estimate how a variable (e.g., the frequency of the word *gun*—word count per time point) changes as a function of time. The instantaneous change in one variable with respect to another variable is known as a *derivative*. The derivative can represent different aspects of change. The first derivative of a word frequency's time series, for example, estimates the rate of change of the word or the velocity at which a word's frequency is changing over time. A negative first order derivative indicates that a word is being used less and less, while a positive first order derivative indicates that a word is being used more often as a function of time. The second derivative indicates the speed of the rate of change or the speed of how quickly a word's frequency is changing (i.e., acceleration). A positive second order derivative indicates an "acceleration" in the rate of change of a word's frequency, while a negative second order derivative indicates a deceleration.

In the GLLA framework (Boker et al., 2010; Deboeck et al., 2009), the derivatives are estimated as:

$$\mathbf{Y} = \mathbf{XL}(\mathbf{L}'\mathbf{L})^{-1} \tag{4}$$

where $\mathbf{Y}$ is a matrix of derivative estimates, $\mathbf{X}$ is a time delay embedding matrix (with n embedding dimensions; to simplify the notation, $\mathbf{X} = \mathbf{X}^{(n)}$), and $\mathbf{L}$ is a matrix with the weights expressing the relationship between the embedding matrix and the derivative estimates. The weight matrix $\mathbf{L}$ is a $n \times \alpha$ matrix, where n is the number of embedding dimensions and α is the (maximum) order of the derivative. Each column of the weight matrix is estimated as follows, considering the order of the derivatives going from zero to k, $\alpha = [0, 1, ..., k]$:

$$\mathbf{L}_\alpha = \frac{[\Delta_t(v - \bar{v})]^\alpha}{\alpha!} \tag{5}$$

where Δ_t is the time between successive observations in the time series, v is a vector from one to the number of embedded dimensions (i.e., $v = [1, 2, ...n]$), $\bar{v}$ is the mean of v, α is the order of the derivative of interest, and $\alpha!$ is the factorial of α.

Continuing our example in which W_1 is a time series of the word *gun*'s frequency from time $t = 1$ to $t = 10$, considering a time delay embedding matrix with five dimensions, derivatives up to the second order (i.e., $\alpha = [0, 1, 2]$), and a Δ_t of one, the weight matrix $\mathbf{L}$ is:

$$\mathbf{L} = \left[\frac{[1(v-\bar{v})]^0}{0!}, \frac{[1(v-\bar{v})]^1}{1!}, \frac{[1(v-\bar{v})]^2}{2!}\right] = \begin{bmatrix} 1 & -2 & 2.0 \\ 1 & -1 & 0.5 \\ 1 & 0 & 0.0 \\ 1 & 1 & 0.5 \\ 1 & 2 & 2.0 \end{bmatrix} \tag{6}$$

Applying Equation 4 to estimate the derivatives, $\mathbf{Y}$ is:

$$\mathbf{Y} = \begin{bmatrix} 6.5 & 1 & 0 \\ 7.5 & 1 & 0 \\ 8.5 & 1 & 0 \\ 9.5 & 1 & 0 \\ 10.5 & 1 & 0 \\ 11.5 & 1 & 0 \\ 12.5 & 1 & 0 \end{bmatrix} \tag{7}$$

where the first, second and third column represents the zeroth (observed values), first (rate of change), and second derivative (speed of the rate of change), respectively.

The process described above is repeated for each variable (e.g., each time series of word counts), and then the resulting derivatives can be column bound to form a matrix $\mathbf{D}$ for each individual (e.g. each Twitter account). By following this process, both linear and nonlinear dynamics are preserved for each individual. Different levels of analysis can then be implemented. If the goal is to investigate the *population* structure, then the $\mathbf{D}$ matrices can be stacked and EGA can be used to estimate the number of underlying dimensions using data from all individuals. If the data contains multiple groups of individuals, for example right- and left-leaning trolls, then the the $\mathbf{D}$ matrices can be stacked by group, and EGA is applied separately in each resulting stacked matrix, generating one dimensionality estimation per group. Finally, if the goal of the analysis is in the intraindividual structure, then EGA can be used in each $\mathbf{D}$. The result, in this case, will be a structure for each individual, separately. Irrespective of the level of analysis (population, group, or individual), the resulting clusters in the network corresponds to variables (words) that are changing together. This is, in summary, the general idea behind our Dynamic Exploratory Graph Analysis (DynEGA) approach.

How DynEGA can be used to extract the latent trends (or topic/factor scores) for each topic/factor

Network Loadings

In these topics, certain terms contribute more information to topical trends than others. Measures to quantify the contribution of information at the nodal-level (i.e., term-level) are called centrality measures. Centrality measures quantify the relative position of terms based on their connections to other nodes in the network. One of the most common measures is called *node strength*, which corresponds to the sum of a term's connections in the network. In a series of simulation studies, Hallquist, Wright, and Molenaar (2019) demonstrated that node strength was roughly redundant with confirmatory factor analysis (CFA) loadings. They found, however, that a node's strength represents a combination of dominant and cross-factor loadings.

Considering this limitation of node strength, a more recent simulation study evaluated node strength when it's split by dimensions (or topics) in the network (Christensen & Golino, 2020). In this study, Christensen and Golino (2020) mathematically defined a measure called *network loadings* by splitting a node's strength based on its connections within and between dimensions identified by EGA. This measure was then standardized to derive an equivalent measure to factor loadings. Their simulation study demonstrated that network loadings can effectively estimate the population (or true) loadings and are roughly equivalent with exploratory factor analysis (EFA) loadings. Notably, like CFA loadings, network loadings had zeros in the loading matrix from nodes in the network that were not connected. This places network loadings on a middle ground between a saturated (EFA) and simple structure (CFA). Below, we provide mathematical notation for how network loadings are computed.

Let W represent a symmetric $m \times m$ matrix where m is the number of terms. Node strength is then defined as:

$$NS_i = \sum_{j}^{m} w_{ij}, \tag{8}$$

where w_{ij} is the weight (e.g., partial correlation) between node i and node j, and NS_i is the sum of the weights between node i and all other nodes. Using the definition of node strength (8), we can define node strength split between the communities, C, identified by EGA:

$$NL_{ic} = \sum_{j \in c}^{C} w_{ij}, \tag{9}$$

where w_{ij} is the weight of node i with the subset of nodes j that belong to community c (i.e., $j \in c$), and NL_{ic} is the sum of the weights for node i in community c (or unstandardized network loading for node i in community c). From the unstandardized network loadings (9), standardized network loadings follow with:

$$z_{NL_{ic}} = \frac{NL_{ic}}{\sqrt{\sum_{c}^{C} NL_c}}, \tag{10}$$

where NL_c is the sum of network loadings in community c and $z_{NL_{ic}}$ is the standardized network loading of node i in community c. It's important to emphasize that network loadings are in the unit of association—that is, if the network consists of partial correlations, then the standardized network loadings are the partial correlation between each node and dimension.

Network Scores

Importantly, these network loadings form the foundation for computing network scores, which can be used to extract information about the topics in the network. Because the network loadings represent the middle ground between a saturated (EFA) and simple (CFA) structure, the network scores accommodate the inclusion of only the most important cross-loadings in their computation. This capitalizes on information often lost in typical CFA structures but reduces the cross-loadings of EFA structures to only the most important loadings. Below we detail the mathematical notation for computing network scores.

First, we take each community and identify items that do not have loadings on that community equal to zero:

$$z_{tc} = z_{NL_{i \in c}} \neq 0, \tag{11}$$

where z_{NL_c} is the standardized network loadings for community c, and z_{tc} is the network loadings in community c,that are not equal to zero. Next, z_{tc} is divided by the standard deviation of the corresponding items in the data, X:

$$wei_{tc} = \frac{z_{tc}}{\sqrt{\frac{\sum_{i=1}^{t \in c}(X_i - \bar{X})^2}{n-1}}}, \tag{12}$$

where the denominator, $\sqrt{\frac{\sum_{i=1}^{t \in c}(X_i - \bar{X}_t)^2}{n-1}}$, corresponds to the standard deviation of the items with non-zero network loadings in community c, and wei_{tc} is the weight for the non-zero loadings in community c. These can be further transformed into relative weights for each non-zero loading:

$$relWei_{tc} = \frac{wei_{t \in c}}{\sum_{c}^{C} wei_{t \in c}}, \tag{13}$$

where $\sum_{c}^{C} wei_{t \in c}$ is the sum of the weights in community c, and $relWei_{tc}$ is the relative weights for non-zero loadings in community c. We then take these relative weights and multiply them to the corresponding items in the data, $X_{t \in c}$, to obtain the community (i.e., topic) score:

$$\hat{\theta}_c = \sum_{c}^{C} X_{t \in c} \times relWei_{t \in c}, \tag{14}$$

where $\hat{\theta}_c$ is the network score for community c.

It is interesting to point that one of the first researchers to discover the equivalence between network models and factor models was Guttman (1953). Although network was not yet a specific area of research, especially in Psychology, Guttman (1953) proposed a new psychometric technique termed *"image structural analysis"* which is basically a network model with node-wise estimation using multiple regression. Guttman demonstrated how his new psychometric technique relates to factor models, and pointed that the former is a special case of the node-wise network model where the errors of the variables are orthogonalized. Therefore, it should be expected that the network scores we define should be directly related to traditional factor scores.

Checking the feasibility of Dynamic Exploratory Graph Analysis to estimate the number of underlying factors/topics

Now that the DynEGA technique was described in detail, an important question must be addressed: how accurate is DynEGA to estimate the number of latent topics (or latent factors). This section will briefly address this question.

A plausible underlying mechanism of latent topics can be represented as a direct autoregressive factor score model (**DAFS**; Engle & Watson, 1981), which is characterized by the autoregressive structure of the latent dimensions (Nesselroade et al., 2002). Since our paper focuses on modeling text data, we will adjust the nomenclature accordingly. In the **DAFS** framework, the observed variables $\mathbf{w}_t$ at time t ($t = 1, 2, ..., N$) are given by:

$$\mathbf{w}_t = \boldsymbol{\Lambda}\mathbf{f}_t + \mathbf{e}_t \tag{15}$$

where $\mathbf{\Lambda}$ is the topic (factor) loading matrix (a $p \times q$ matrix), $\mathbf{f}_t$ is a $q \times 1$ vector of topics at time t, and $\mathbf{e}_t$ is a $p \times 1$ vector with measurement errors following a multivariate normal distribution with mean zeros and covariance matrix Q (Nesselroade et al., 2002; Zhang, Hamaker, & Nesselroade, 2008).

The topic scores, $\mathbf{f}_t$, are given by:

$$\mathbf{f}_t = \sum_{l=1}^{L} \mathbf{B}_l \mathbf{f}_{t-l} + \mathbf{v}_t \tag{16}$$

where $\mathbf{B}_l$ is a $q \times q$ matrix of autoregressive and cross-regressive coefficients, $\mathbf{f}_{t-l}$ is a vector of topic score l occasions prior to occasion t and $\mathbf{v}_t$ is a random shock vector (or innovation vector) following a multivariate normal distribution with mean zeros and $q \times q$ covariance matrix $\mathbf{D}$ (Nesselroade et al., 2002; Zhang et al., 2008). In the *DAFS* model, $\mathbf{\Lambda}$, $\mathbf{B}_l$, $\mathbf{Q}$ and $\mathbf{D}$ are invariant over time.

Data following the *DAFS* model can be simulated using the `simDFM` function of the *EGAnet* package (Golino & Christensen, 2019). Below we present a brief simulation investigating how accurate is DynEGA to recover the number of simulated topics. We also investigate the distribution of variables per topic and the correlation between the simulated and the estimated topic scores. Accuracy can be calculated as follows:

$$Acc = \frac{\Sigma_C}{N}, \text{ for } C = \begin{cases} 1 \text{ if } \hat{\theta} = \theta \\ 0 \text{ if } \hat{\theta} \neq \theta \end{cases} \tag{17}$$

Where $\hat{\theta}$ is the estimated number of latent topics, θ is the true number of latent topics used to simulate the data (i.e. *ground truth*), and N is the number of sample data simulated.

The distribution of the variables per topic can be checked using *normalized mutual information* (NMI; Horibe, 1985). NMI is used to compare the similarity between two vectors (of discrete variables) and assigns a value of zero where the two vectors are totally dissimilar, and a value of one where they are identical in an information theoretic perspective. For example, consider two vectors (v_1 and v_2) representing the partition of a multidimensional space into two groups, so that $v_1 = (1,1,1,1,1,2,2,2,2,2)$ and $v_2 = (2,2,2,2,2,1,1,1,1,1)$. The NMI of the two vectors equals one, since both vectors are presenting the same information (i.e., that the multidimensional space is grouped in two dimensions composed by the first three elements and the last three elements, respectively). If v_1 or v_2 were to be compared to a third vector $v_3 = (1,1,1,2,2,2,2,3,3,3)$, the NMI between v_1 or v_2 and v_3 equals 0.38. Clearly, v_1 and v_2 exhibits the same partitioning of the multidimensional space, different from the partitioning of v_3.

The two vectors used to compute NMI are the vector of the *assigned variables per topic*—that generates the block diagonal matrix $\mathbf{\Lambda}$—equation 15, and the vector containing the estimated topic number per variable.

Simulation Design

In this brief simulation, five relevant variables were systematically manipulated using Monte Carlo methods: the number of time points (i.e., the length of the time series), topic (factor) loadings, number of variables per topic, measurement error and type of observed data. For each of these, their levels were chosen to represent conditions that are encountered in dimensionality assessment simulation studies (e.g., Garrido, Abad, & Ponsoda, 2013; Golino et al., 2020; Zhang et al., 2008) and that could produce differential levels of accuracy for DynEGA and the LDA techniques. It is important to point that most studies in the area of topic modeling do not implement Monte-Carlo simulations to test the techniques and algorithms developed to estimate the number of latent topics (see: Arun et al., 2010; Blei et al., 2003; Deveaud et al., 2014), using what is called *empirical evaluation* of LDA. In other words, the authors develop a new topic modeling technique (or an algorithm to decide the optimal n number of topics to be extracted by LDA) and apply it to real-world text data. In the present paper, we decided to implement a brief simulation, so we could investigate how reliable is DynEGA to estimate the number of latent topics, and also to compare this new technique to the widely used LDA. Therefore, the levels of the variables systematically manipulated in the current simulation were decided based on studied from the area of dimensionality assessment rather then from the topic modeling literature.

For the *length of the time series*, three conditions were used: 50, 100, and 200. The number of time points were selected based on the conditions tested by Zhang et al. (2008).

Topic (factor) loadings were simulated with the levels of .40, .55, .70, and 1. According to Comrey and Lee (2016), loadings of .40, .55, and .70 can be considered as poor, good, and excellent, respectively, thus representing a wide range of factor saturations. In addition, loadings of 1 were also simulated to allow for the evaluation of the DynEGA technique under ideal conditions.

Number of variables per topic were composed of 5 and 10 indicators. In the dimensionality assessment literature there is a consensus that three variables are the minimum required for factor identification (Anderson, 1958). In the present simulation 5 items per topic represents a slightly overidentified model, while latent structures composed of 10 variables may be considered as highly overidentified (Velicer, 1976; Widaman, 1993).

The *measurement error* covariance matrix had two conditions (i.e. two diagonal matrices), one with 0.15^2 and the other with 0.25^2 in the diagonal. The choice to use small measurement errors was to verify how the methods perform under minimum error conditions, so the impact of the other variables systematically manipulated in our simulation could be better understood.

Two *types of observed data*, normal continuous and ordered categorical, generated. For the ordered categorical data, a function to categorize the data based on Garrido et al. (2013) and Golino et al. (2020) was used. First, the normal continuous data was simulated, and then the values of the simulated observed variables were discretized to four categories. The LDA techniques were applied only to the ordered categorical data, since they cannot handle normal continuous variables.

Three variables were held constant: (a) the number of topics (three), (b) the matrix with autoregressive (0.8) and cross-regressive coefficients (0), and (c) the covariance matrix for the random shock (off-diagonal = 0.18; diagonal = 0.36). The values of the autoregressive and cross-regressive coefficients and the random shock matrix were selected following Zhang et al. (2008), to ensure the topics were stationary time series (although the DynEGA model can also model non-stationary time series).

Data Generation

For each combination of variables systematically controlled in the Monte-Carlo simulation, 500 data matrices were generated according to the *DAFS* model (equation 15). Given the predefined parameter values, each data matrix was generated as follows. First, the matrix of random shock vectors $\mathbf{v}_t$ is generated following a multivariate normal distribution with mean zeros and 3×3 covariance matrix $\mathbf{D}$ (off-diagonal values = 0.18; diagonal values = 0.36), where t is the number of time points plus 1,000 (used as the burn-in estimates for the chain). Second, the topic (factor) scores are calculated using equation 16 and the first 1,000 estimates are removed (burn-in phase). Third, the measurement error matrix is estimated following a multivariate normal distribution with mean zeros and $p \times p$ covariance matrix Q, where p is the total number of variables (number of variables per topic times three). Finally, the observed variables are calculated using equation 15. Each data matrix represents data from single individuals.

Data Analysis

The DynEGA technique was implemented using the `EGAnet` package (Golino & Christensen, 2019), and the following arguments of the `dynEGA` function were used. The number of embedding dimensions was set to five, τ (time-lag) was one and the time between successive observations for each time series was one (delta). Two network methods (*glasso* and *TMFG*) were used to construct the networks of the first derivatives computed using the GLLA model. The correlation between the first order derivatives was computed using Pearson's coefficient and were used as the input to estimate the networks.

The LDA technique was implemented via the `topicmodels` package (Hornik & Grun, 2011) using a Gibbs sampling estimator (Phan, Nguyen, & Horiguchi, 2008). Readers interested in specific details of the LDA estimation methods are referred to Hornik and Grun (2011). Since LDA requires the number of topics to be specified by the user, five approaches to estimate the optimal number of latent topics were used. The first two estimate from two to six topics and calculate the AIC and BIC of the resulting LDA solution. The remaining four approaches estimate from two to six topics and select the number of topics using the algorithms developed by Arun et al. (2010) (**Arun**), Cao et al. (2009) (**Cao**) and Deveaud et al. (2014) (**Deveaud**) using the `ldatuning` package (Nikita, 2016).

The algorithm developed by Arun et al. (2010) selects the number of latent topics that minimize the Kullback-Liebler divergence between the matrix representing word probabilities for each topic and the topic distribution within the corpus (Hou-Liu, 2018). The algorithm developed by Cao et al. (2009) selects the number of topics based on topic density, searching for the number of topics that minimizes the average cosine similarity between topic

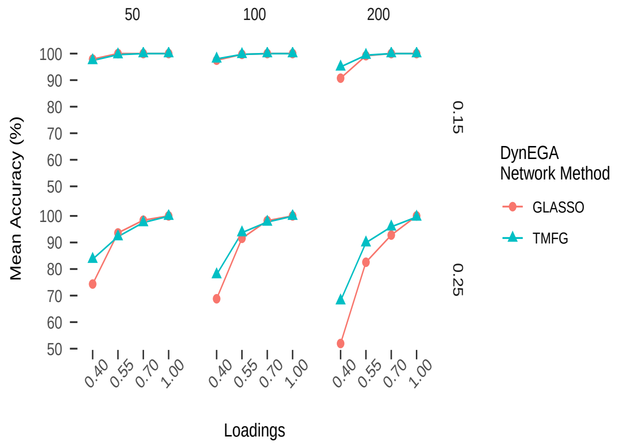

Figure 1. Mean accuracy per network method used in the DynEGA technique, magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

distributions, while Deveaud et al. (2014) developed an algorithm that selects the optimum number of topics by maximizing the average Jensen-Shannon distance between all pairs of topic distributions (Hou-Liu, 2018). LDA was applied in the ordered categorical data condition only, since it cannot handle continuous variables. All code used in the current paper are available in an online repository at the Open Science Framework platform for reproducibility purposes (see: https://osf.io/4ya6x/?view_only=b6078b404e3049818b359ae0d514f966).

Results

Continuous data

In the continuous data condition, the mean accuracy of DynEGA using *glasso* (ACC_{glasso} = 93.17%) and *TMFG* (ACC_{TMFG} = 95.78%) were very similar, as were the mean normalized mutual information (NMI_{glasso} = 96.62%, NMI_{TMFG} = 95.38%). Figures 1 and 2 show that the mean accuracy and normalized mutual information increase with the magnitude of the loadings, but decrease with the increase in the measurement error. It is interesting to note that although the *TMFG* network method is more accurate (see Figure 1), the *glasso* approach gives the higher normalized mutual information, suggesting that the latter method more accurately allocate the variables into the correct latent topics. Both figures also show that the mean accuracy and normalized mutual information decreases with the increase in the measurement error.

Figure 3 shows the mean correlation between simulated and estimated topic scores per network method used in the DynEGA technique, magnitude of the loadings, number of time points and magnitude of the measurement error. The DynEGA technique with the *glasso* network method presented a higher mean correlation (GLASSO) than the *TMFG* network method (TMFG). The differences are higher for factor loadings of 0.40 (see Figure 3).

Ordered categorical data

In the ordered categorical data condition, the scenario is slightly different. The mean accuracy of DynEGA using the *TMFG* network method (ACC_{TMFG} = 85.70%) was

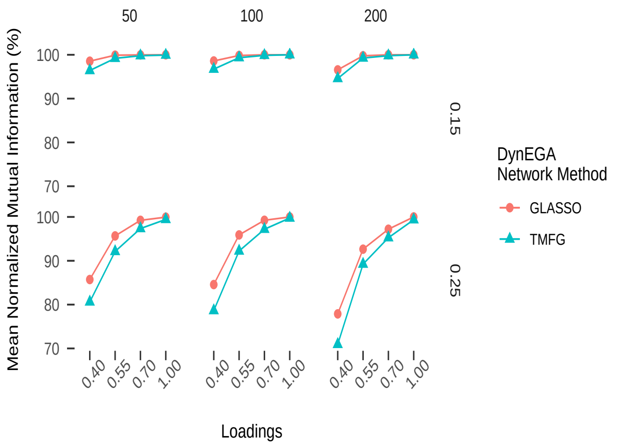

Figure 2. Mean normalized mutual information per network method used in the DynEGA technique, magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

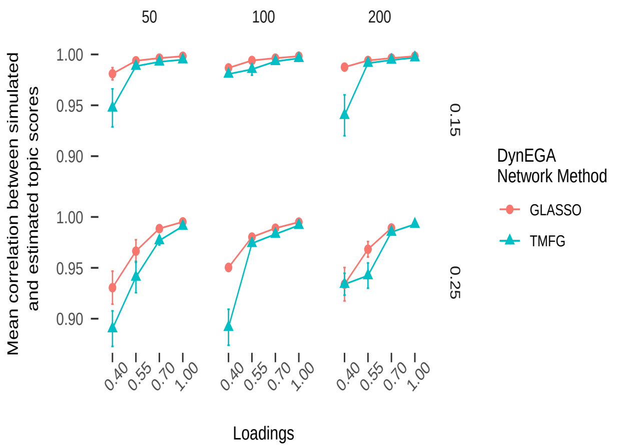

Figure 3. Mean correlation between simulated and estimated topic scores per network method used in the DynEGA technique, magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

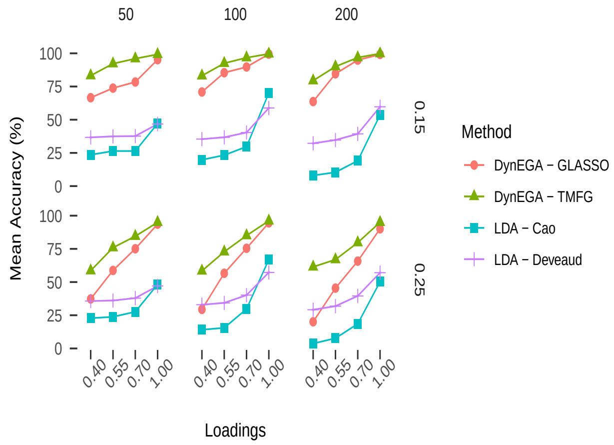

Figure 4. Mean accuracy per method in the ordered categorical data condition. Magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

higher than the mean accuracy of *glasso* (ACC_{glasso} = 73.43%). As with the continuous data type, DynEGA with the *glasso* network method presented a higher mean normalized mutual information than DynEGA with *TMFG* (NMI_{glasso} = 88.20%, NMI_{TMFG} = 85.67%).

The strategies used to select the optimal number of topics via LDA all presented a very low accuracy. The *Arun* algorithm (Arun et al., 2010) had an accuracy of zero, since it selected the maximum number of topics compared (six) as the optimal number of topics all the time. AIC and BIC presented a mean accuracy of 14.62% and 16.84%, respectively. *Cao's* (Cao et al., 2009) and *Deveaud*'s algorithm (Deveaud et al., 2014) presented the best mean accuracy: 28.60% and 31.24%, respectively. Figure 3 shows the mean accuracy per method in the ordered categorical data condition, per loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets). The DynEGA technique with the *TMFG* network method has the highest accuracy, followed by DynEGA with the *glasso* network method, the *Deveaud*'s algorithm and *Cao*'s algorithm for LDA.

Figure 5 shows that the DynEGA with the *glasso* network method presented a slightly higher mean normalized mutual information than the DynEGA with the *TMFG* network method, specially with loadings of 0.40, 0.55 and 0.70.

Figure 6 shows the mean correlation between simulated and estimated topic scores per network method used in the DynEGA technique, magnitude of the loadings, number of time points and magnitude of the measurement error. The DynEGA technique with the *glasso* network method presented a higher mean correlation between the simulated and estimated topic scores (GLASSO) than the *TMFG* network method (TMFG). The differences are higher for loadings of 0.40 (see Figure 6).

Applying the dynamic exploratory graph analysis in the IRA-linked Twitter data.

The substantive problem this paper addresses is related to the use of social media, especially Twitter, by agencies or groups of people devoted to exploit social divisions in

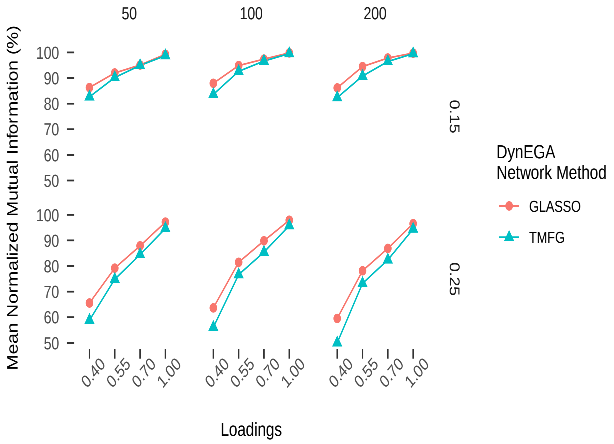

Figure 5. Mean normalized mutual information per network method used in the DynEGA technique, magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

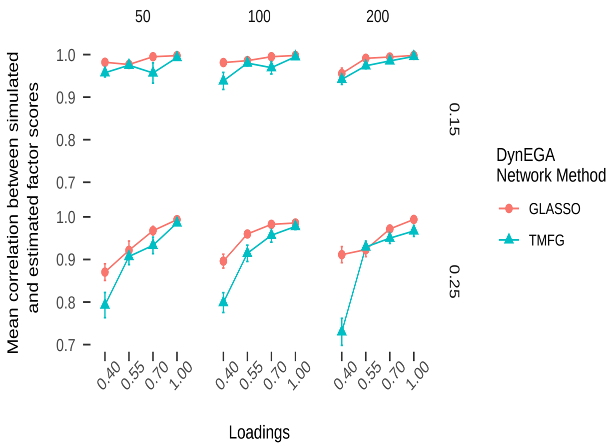

Figure 6. Mean correlation between simulated and estimated topic scores for the ordered categorical data condition per network method used in the DynEGA technique, magnitude of the loadings (x-axis), number of time points (vertical facets) and magnitude of measurement error (horizontal facets).

a society for political reasons. People are increasingly using online social platforms to communicate their ideas and politicians and government bodies are taking advantage of these platforms to leverage their interaction with voters/citizens. Social media platforms are becoming more and more central in debates about relevant issues such as abortion, gun control, and other controversial topics (Rajadesingan & Liu, 2014), sometimes leading to strongly polarized positions (Yardi & Boyd, 2010).

In the U.S., online activities were used to influence public opinion and voters during the 2016 campaign, with attacks occurring before and during the electoral process (Linvill & Warren, 2018). Twitter accounts linked to the Internet Research Agency (IRA), based in Russia, were used to try to divide voters in a wide range of issues (Linvill & Warren, 2018), and their information was released by the US congress after an investigation to discover Russian state-sponsored trolls (Zannettou, Caulfield, De Cristofaro, et al., 2019). In this section the DynEGA technique is applied to a large database of IRA-linked Twitter accounts extracted by Linvill and Warren (2018), and posted online by the FiveThirtyEight team (Roeder, 2018). The goal is to investigate the strategies used by right- and left-wing Twitter accounts.

The original data contains almost 3 million Twitter posts, by 2843 unique accounts, starting in January, 2013 to May, 2018. Linvill and Warren (2018) classified the accounts into five groups: right troll, left troll, news feed, hashtag gamer, and fearmonger. The first two types of accounts mimic right or left-leaning people. The news feed accounts present themselves as local news aggregators, the hashtag gamer accounts specialize in playing hashtag games, and the fearmonger accounts spread news about a fake crisis (Roeder, 2018). In the current analysis, only accounts identified as right- and left-leaning trolls are used, with posts (not including retweets) from January 2016 to January 2017. Accounts with less than 50 posts were excluded, resulting in 276,752 posts and 236 accounts. The 236 accounts included in the analysis can be considered influential, since they have a non-negligible number of followers (2,014 on average). The high number of followers can help the trolls in pushing specific narratives to a much greater number of Twitter users (Zannettou, Caulfield, De Cristofaro, et al., 2019).

The twitter posts were split by account type (i.e. right and left trolls) and pre-processed using the `tm` package from R (Feinerer, Hornik, & Meyer, 2008). URLs were removed from the text data as well as punctuation, numbers, and stop words. All characters were converted to lowercase and the words were stemmed (i.e., reduced to their stem, base or root, using *Porter's algorithm*; Porter, 1980). The sparsity of the resulting document term matrix (i.e. a data frame where the columns are unique words and rows are different documents or posts) was decreased using a threshold of 0.99 for the right trolls and 0.993 for the left trolls, so that words with a sparsity above the threshold is removed from the document term matrix, resulting in 108 unique words for the right and 113 words for the left trolls. In text mining, deciding the number of words to use is done arbitrarily. In the current analysis, we decided to use around 100 words per account type, so the final document term matrix would have enough words to capture different types of topics, but not very specific topics (or niche content), which could happen if the number of words used increases (since the number of words used depends on the sparsity threshold, using more words means using words that are less frequently used).

To estimate the topics per account type, the DynEGA technique was used via the `dynEGA` function from the `EGAnet` package (Golino & Christensen, 2019). Since the *TMFG* network method presented the highest accuracy in the estimation of the number of topics in the ordered categorical data, it was the network construction method used in the current analysis. The `dynEGA` arguments were set as follows: the number of embedding dimensions was set to five, τ (time-lag) was set to one and the time between successive observations for each time series was one (delta). The correlation between the first order derivatives was computed using Pearson's coefficient and were used as the input to estimate the network. The level of analysis used in this section was set to *population*, meaning that the derivatives are estimated per Twitter account and then row binded, creating a long data frame that is used to compute the correlation matrix. The estimated topics are, then, the "mean" structure of population.

Figure 7 shows the network structure of the right-leaning trolls. The nodes represent the words and the edges are the Pearson correlation of the words' first order derivatives. Eight topics were estimated. The first topic (red nodes) referred to issues related to job/work in America, and contain words such as *make*, *need*, *work*, and *america*. The second topic (orange nodes) involved words as *Trump*, *win*, *trumpforpresident*, *maga* (the short for *make america great again*, Trump's campaign motto), and is clearly related to supporting Trump for president of the US. The third topic (green nodes) referred to the liberal media (in particular the CNN news network), and contain words such as *show*, *video*, *liber* (root of *liberal*), and *cnn*. The fourth topic focused on terrorism and anti-Islam content, with words such as *terrorist*, *stopislam*, *attack*, and *isi* (a reference to the ISIS terrorist group). The fifth topic (light blue nodes) focused on support for police and the white people movement in the US, with words such as *live*, *support*, *polic* (root of police), and *white*. The sixth topic (dark blue nodes) was the most difficult to interpret, since it contain words related to time (e.g. time, day, one, year). The seventh topic (purple nodes) focused on attacking Hillary Clinton and her campaign, with words as *campaign*, *clinton*, *lie*, *hillaryclinton*, and *email*. Finally, the eight topic (pink nodes) was related to gun control, the *Wake up America* movement (a movement against the suppression of individual rights such as gun ownership), and Ted Cruz (the US Senator from Texas).

To understand how each topic changes over time, the topic scores were estimated using equation 14. Figure 8 shows the mean topic score per date for the right trolls. Some trends are interesting to mention. For example, Topic 2 (*supporting Trump for president of the US*) peaked twice, once in April 27th (day that Trump won in five states in the primary election for the republican party, and day he gave a controversial speech on foreign policy) and once in November 8th, 2016 (election day). Topic 4 (*Terrorism/Attacking Islamism*) also presented two peaks, on March 22nd (day of the bombings at Brussels airport and a metro station that killed 32 people) and on November 28th (day of the Ohio State University terrorist attack). Topic 7 (*Attacking Hillary Clinton*) presented clear peak on August 22 and 27, 2016, when it was released by the media that the FBI found that the Democratic candidate had received or sent 15,000 e-mails from her personal e-mail account while acting as a secretary of state. The examples for topics two, four and seven are presented in Figure 9.

The left-leaning trolls presented a different story. Figure 10 shows the network structure of the left trolls, with nine topics. The first topic (red nodes) probably referred to an online movement generally called *life comes back*, supporting the return of Lamar Odom to the basketball courts after years treating for substance abuse, and include words such as *first*, *day*, *life*, *come*, and *back*. The second topic (orange nodes) involved words as *blacklivesmatter*, *racist*, *make*, *america*, *racism*, and is clearly related to the Black Lives Matter movement, anti-racism activities and against the *maga* movement supporting Trump for president. The third topic (light green nodes) included words such as *really*, *say*, *think*, *like*, *people*, and seems to be encouraging people to say and think what they want. The fourth topic (green) focused live streaming of music and video content related to hip-hop, with words such as *hiphop*, *music*, *nowplaying*, and *watch*. The fifth topic (green/blue-ish nodes) focused the use of Twitter, with words such as *twitter*, *love*, and *media*. The sixth topic (light blue nodes) was related to the police actions, with words as *white*, *police*, *cops*, and *woman*. The seventh topic (dark blue nodes) was the most difficult to interpret, since it contain only three words (years, still and today). The eight topic (purple nodes) was related to the use of Twitter for pro-Black culture, with words as *blacktwitter*, *good*, and *god*. Finally, the ninth topic (pink nodes) focused on the elections, with words as *hillary*, *trump*, and *vote*.

To understand how each topic changes over time, the topic scores were estimated using equation 14. Figure 11 shows the mean topic score per date for the left trolls. Some trends are interesting to mention. For example, Topic 2 (Black Lives Matter movement, anti-racism activities, and against the *maga* movement supporting Trump for president) presented six peaks. Three of them (see Figure 12) coincided with Donald Trump's rallies in Tampa (FL, on 2/12), Gaffney (SC, on 2/18), and St. Louis (MO, 3/11).

Discussion

Quantitative analysis of texts can be very useful for understanding the strategies used in online information warfare. In the past few years a number of studies investigated characteristics of Twitter accounts created to sow discord and polarization before, during, and after the 2016 presidential election in the U.S. (Ghanem et al., 2019; Llewellyn et al., 2018; Zannettou, Caulfield, De Cristofaro, et al., 2019; Zannettou, Caulfield, Setzer, et al., 2019). A common denominator in these studies has been the use of Latent Dirichlet Allocation (LDA; Blei et al., 2003) to estimate latent topics from posts on social media platforms. As pointed out earlier, despite the usefulness of LDA, it has several limitations that make its use in text data from social media platforms doubtful.

In the current paper we introduced a new (and fast) method to estimate latent dimensions (e.g., factors, topics) in

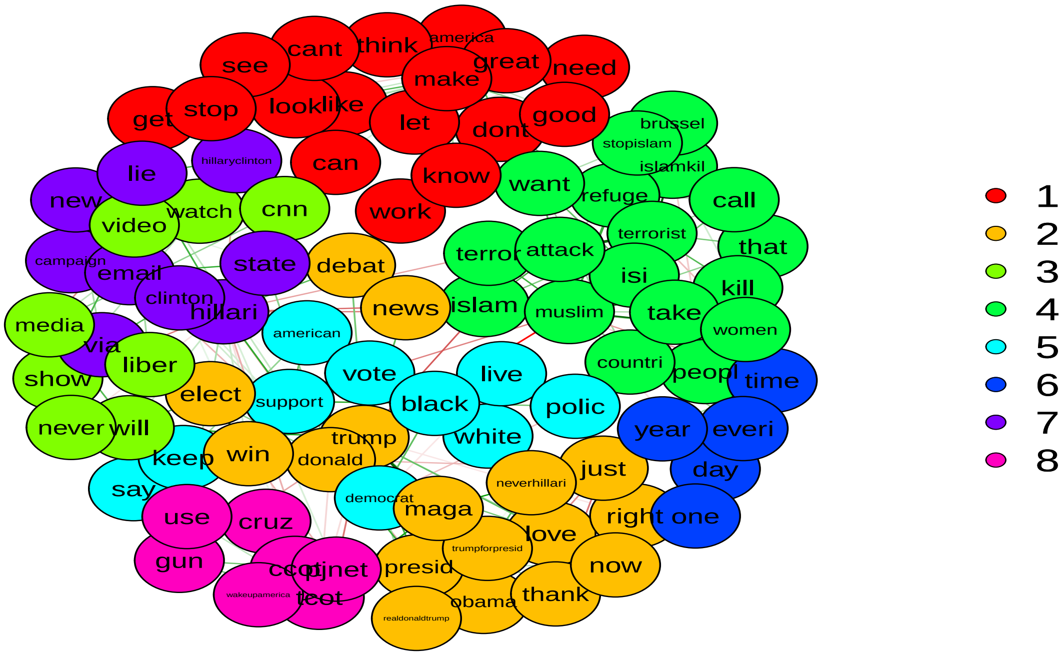

Figure 7. Network structure estimated using DynEGA of the right trolls document term matrix showing eight topics (clusters). The nodes represent the words and the edges are the Pearson correlation of the words' first order derivatives.

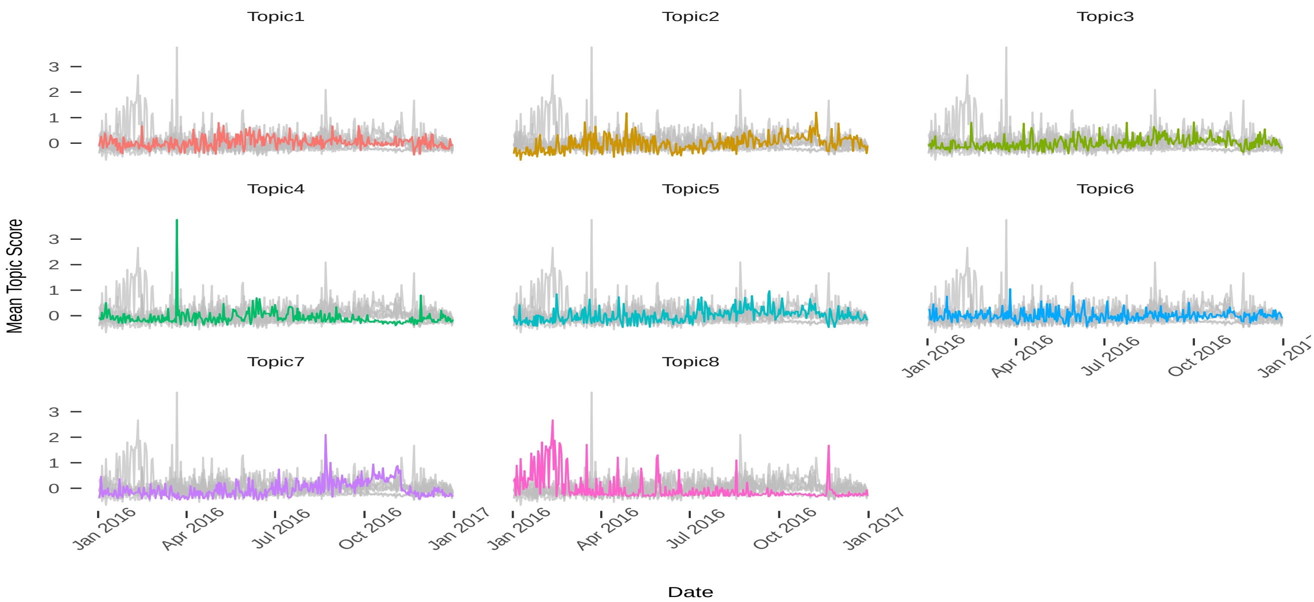

Figure 8. Latent trends of the topics per date.

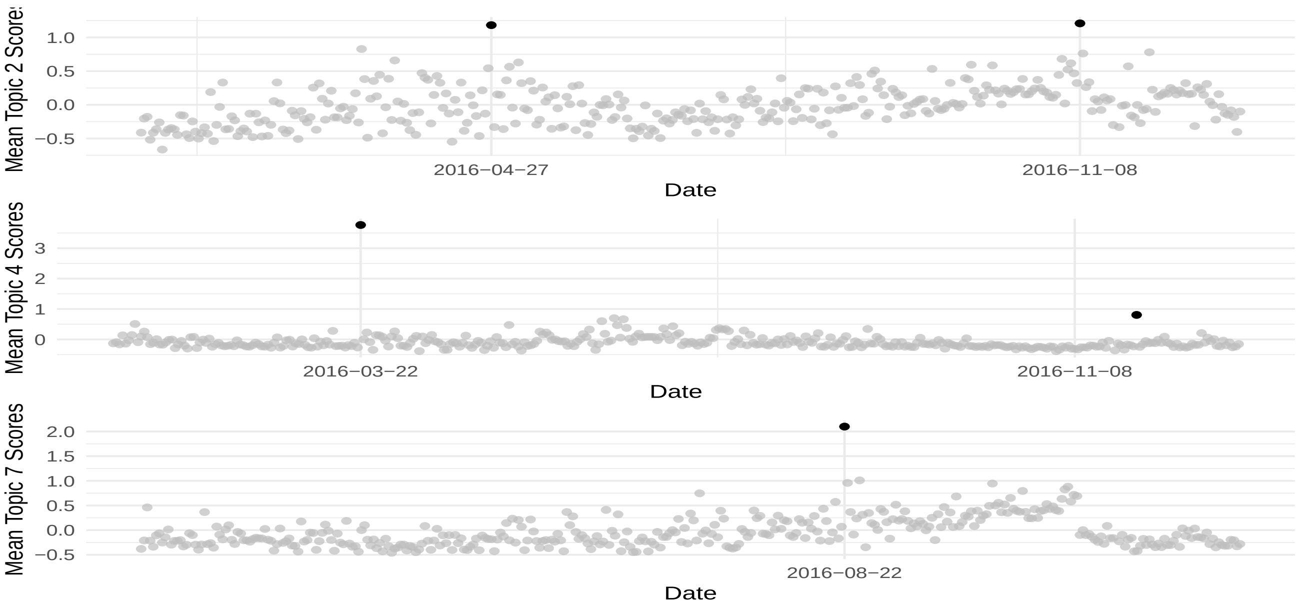

Figure 9. Latent trends of topic two (supporting Trump for US President), topic four (Terrorism/Anti-Islamism), and topic seven (Attacking Hillary Clinton).

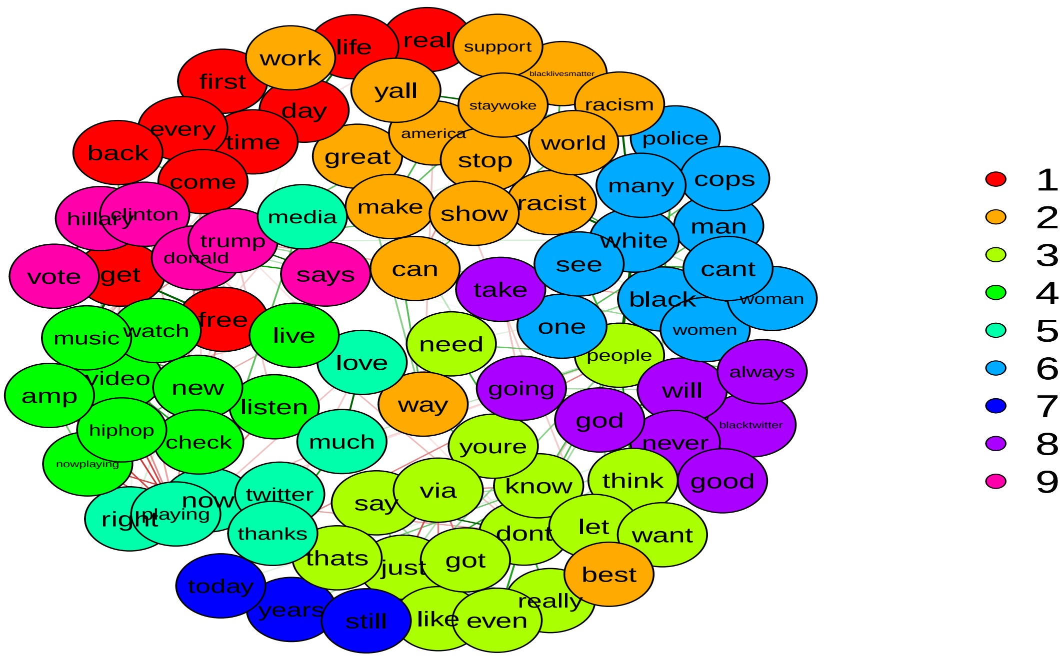

Figure 10. Network structure estimated using DynEGA of the left trolls document term matrix showing nine topics (clusters). The nodes represent the words and the edges are the Pearson correlation of the words' first order derivatives.

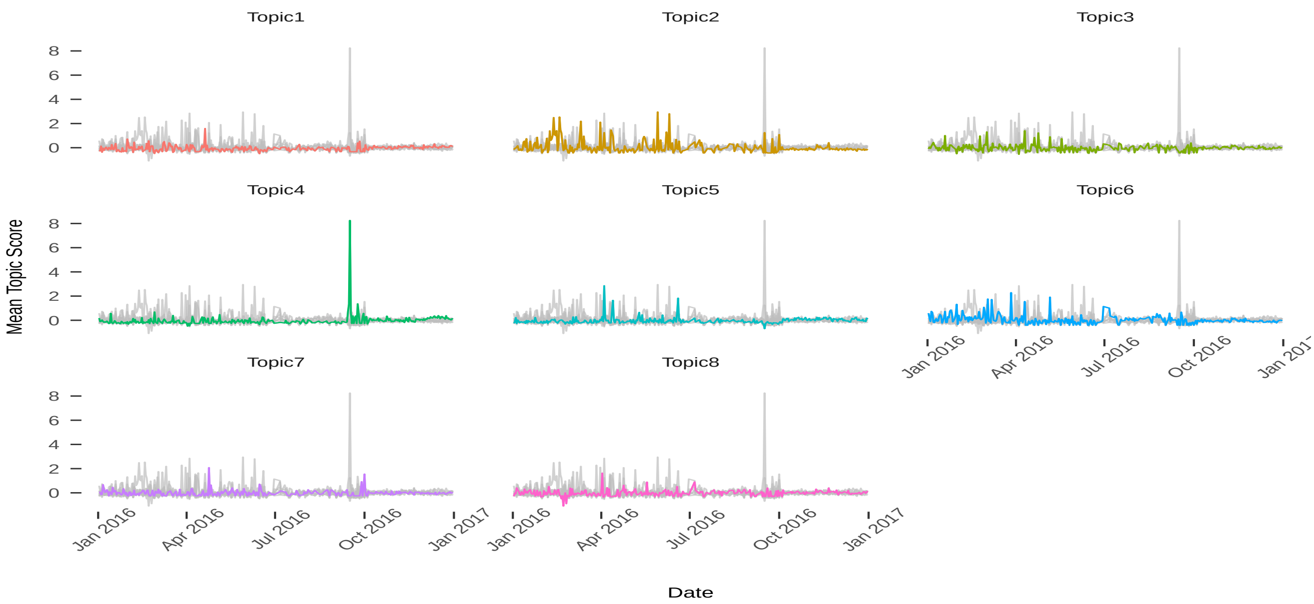

Figure 11. Latent trends of the topics per date - Left Trolls.

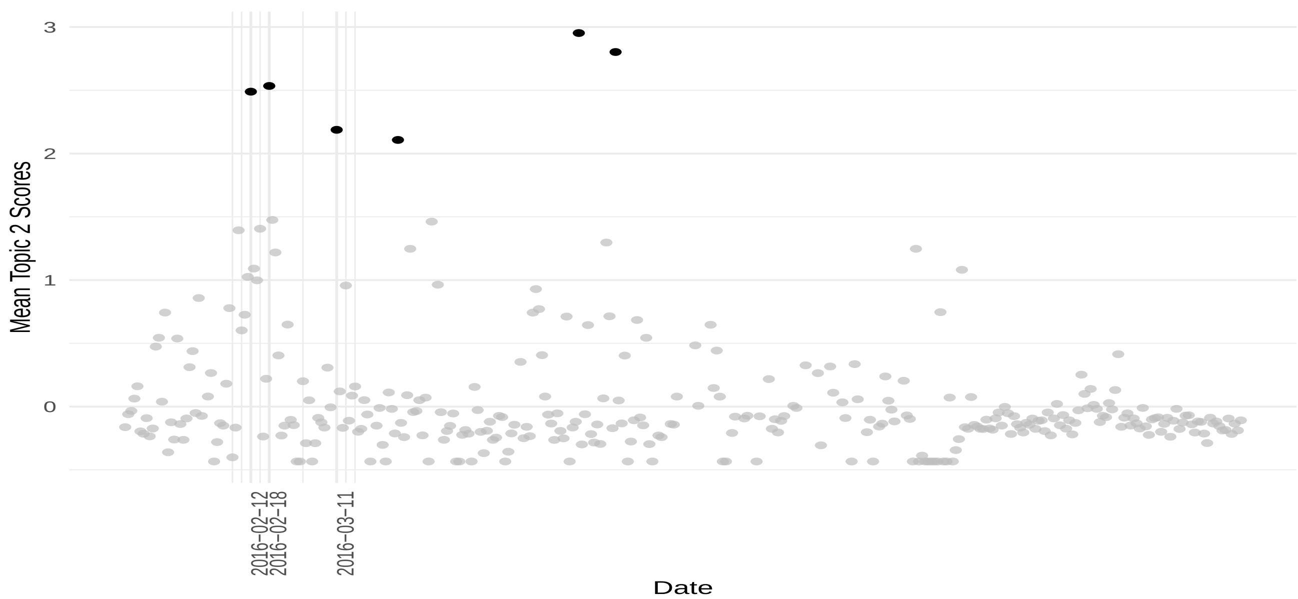

Figure 12. Latent trends of topic two (Black Lives Matter movement, anti-racism activities and against the *maga* movement supporting Trump for president).

multivariate time series data termed Dynamic Exploratory Graph Analysis. The DynEGA technique can be used to estimate the latent structure of topics published in social media (using the time series of word frequencies), improving our capacity to understand the strategy used by accounts created as tools of information warfare. Unlike LDA, DynEGA automatically identifies the number of topics and the distribution of variables (words) per topic and can model temporal dynamics in both stationary and non-stationary time series (although the simulation only focused on stationary time series). A brief Monte-Carlo simulation study was implemented to check the capacity of DynEGA to recover the parameters (i.e. number of topics and topic scores) used to simulate the data using the direct autoregressive factor score model (Engle & Watson, 1981; Nesselroade et al., 2002).

The results showed that the DynEGA technique presented a very high performance in recovering the number of simulated topics, especially when the variables have moderate and high loadings (for both the continuous and ordered categorical data conditions). DynEGA with the *TMFG* network method presented a higher accuracy in estimating the number of simulated topics compared to the *glasso* network method for both types of variables (continuous and ordered categorical), and is the technique we recomend for topic modeling. However, it is important to point that DynEGA with *glasso* presented a slightly higher normalized mutual information and correlation between the simulated and the estimated topic scores.

DynEGA presented an accuracy considerable higher than LDA in the ordered categorical data condition. Of the LDA techniques, *Cao's* (Cao et al., 2009) and *Deveaud*'s algorithms (Deveaud et al., 2014) presented the best mean accuracy among the other LDA techniques, but never exceeded the 75% accuracy threshold (Figure 4). By comparison, DynEGA with the *TMFG* network method was at or above the 75% accuracy threshold in most of the conditions tested (Figure 4). The results of our Monte-Carlo simulation present strong evidence that LDA should be used with caution when applied in data from social media platforms, such as Twitter, if used at all. On the other side, the new technique presented in this paper shows a high accuracy in estimating the number of simulated topics and a high normalized mutual information for the moderate and high loadings condition. Moreover, the DynEGA methods had a very high correlation between the simulated and the estimated topic (factor) scores (Figure 6), irrespective of the condition tested.

The application of DynEGA to a large database of IRA-linked Twitter accounts containing right- and left-leaning trolls (Linvill & Warren, 2018), revealed a very interesting set of online information warfare strategies. The right-leaning trolls were posting content supporting Donald Trump's presidential campaign, defending a political agenda aligned with Trump's *make america great again* movement, posting pro-gun, pro-police, anti-terrorism and anti-Islam content, as well as attacking the Democratic candidate Hillary Clinton. The left-leaning trolls were posting content supporting the Black Lives Matter movement, activism against police brutality, black culture and music, and general terms associated with the election (i.e., Hillary, Trump, vote). Notably, the trolls associated to the Internet Research Agency seemed to follow very closely the news coverage of the 2016 electoral race, and any event that could be used to promote their political agenda.

Using the DynEGA topic score methodology, we discovered that the right-wing trolls were posting more topic-related content than usual in important dates, coninciding with: 1) the day Trump gave a controversial speech on foreign policy, 2) the day of bombings at Brussels airport and a metro station, 3) the day of the Ohio State University terrorist attack, and 4) the day the media released the news that the FBI found that the Democratic candidate had received and sent a huge number of e-mails using her personal account while acting secretary of state.

Rather than only spreading content to advance their political agenda, the right-leaning trolls were following a strategy to amplify important events in the US. This strategy may serve two different goals: disguise themselves as regular citizens and to use daily news events to promote a pre-defined, state-sponsored agenda (Linvill & Warren, 2018; Zannettou, Caulfield, De Cristofaro, et al., 2019). Along the same line, the left-leaning trolls displayed pattern of tweeting spikes that coincided with topic-relevant events specifically related to the Black Lives Matter movement during three of six of Donald Trump's presidential rallies.

It is interesting to point that the type of content pushed by the right- and left-wing trolls differed substantially, in line with previous research (Ghanem et al., 2019; Stewart, Arif, & Starbird, 2018). This indicates that the minds behind the online intelligence (or information warfare) operation targeted contents that could potentially maximize the activation (in terms of online interactions) of a network of Twitter users that shared similar political views, and that could be more prone to engage in more extreme political activities online and offline (Fenton, 2016).

Our application of DynEGA reveals the temporal trends of topics solicited by right- and left-wing trolls before, during, and after the 2016 presidential election in the U.S. This analysis revealed topics that were pertinent to several consequential events in the election cycle, demonstrating the coordinated effort of trolls capitalizing on current events in the U.S. This demonstrates the potential power of our approach for revealing temporally relevant information from qualitative text data. Such an approach has applications that extend far beyond the election cycle. Twitter, for example, contains all sorts of information related to people's preferences and patterns of behavior that may be relevant for other data mining endeavors such as advancing our understanding of idiographic personality (e.g., Bleidorn & Hopwood, 2019). Outside of Twitter, DynEGA opens up opportunities for quantifying daily diaries and essays in a meaningfully and concise way. One example might be extracting the underlying affective information in daily diaries of clinical samples to determine whether certain people are responding to treatment or prone to remission. DynEGA further supports analyses across levels—from individuals to groups to population—which opens up avenues for evaluating the connections between within- and between-person structures, which is a timely topic of interest in a number of areas in psychology (e.g., personality; Baumert et al., 2017).

The dynamic modeling of texts is an increasingly important topic in psychology because of the amount of information that's stored in qualitative data. Whether its over social media or qualitative coding of ecological behaviors, effectively extracting quantitative information that can be leveraged to make predictions about people's behavior is underdeveloped area in research. Here, we provide a tool that can quantify this information in a meaningful way so that researchers can complementary identify manifestions of behavior. DynEGA opens the door for researchers to understand their samples from an idiosyncratic to population level, which enables a holistic perspective of the phenomena they're investigating. The richness of qualitative data has always been appreciated but undervalued in many scientific arenas including psychology. Our approach takes one step towards providing researchers with a tool that can reclaim the utility of text data.

References

Ananiadou, S., & McNaught, J. (2006). *Text mining for biology and biomedicine*. Boston: Artech House Publishers.

Anderson, H., T. W. & Rubin. (1958). Statistical inference in factor analysis. In *Proceedings of the 3rd berkeley symposium on mathematics, statistics, and probability* (Vol. 5, pp. 111–150).

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In R. B. Zaki M. J. Yu J. X. (Ed.), *Advances in knowledge discovery and data mining.* (Vol. 6118, pp. 391–402). Springer, Berlin. https://doi.org/https://doi.org/10.1007/978-3-642-13657-3_43

Barfuss, W., Massara, G. P., Di Matteo, T., & Aste, T. (2016). Parsimonious modeling with information filtering networks. *Physical Review E*, *94*(6), 062306.

Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., ... Wrzus, C. (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality*, *31*, 503–528. https://doi.org/10.1002/per.2115

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. In *Proceedings of the 23rd international conference on machine learning* (pp. 113–120).

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, *1*(1), 17–35.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bleidorn, W., & Hopwood, C. J. (2019). Using machine learning to advance personality assessment and theory. *Personality and Social Psychology Review*, *23*, 190–203. https://doi.org/10.1177/1088868318772990

Boker, S. M. (2018). Longitudinal multivariate psychology. In E. Ferrer, S. M. Boker, & K. J. Grimm (Eds.) (pp. 126–141). Routledge.

Boker, S. M., Deboek, P. R., Edler, C., & Keel, P. (2010). Generalized local linear approximation of derivatives from time series. In S. M. Chow, E. Ferrer, & F. Hsieh (Eds.), *The notre dame series on quantitative methodology. Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 161–178). Routledge/Taylor & Francis Group.

Bork, R. van, Borkulo, C. D. van, Waldorp, L. J., Cramer, A. O., & Borsboom, D. (2018). Network models for clinical psychology. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, *5*, 1–35.

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*(9), 1089–1108.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121.

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive lda model selection. *Neurocomputing*, *72*(7-9), 1775–1781.

Cattell, R. B. (1965). Studies in psychology. In C. Banks & P. L. Broadhurst (Eds.) (pp. 223–266). University of London Press London.

Chaney, A. J. B., & Blei, D. M. (2012). Visualizing topic models. In *Proceedings of the sixth international aaai conference on weblogs and social media*.

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771.

Christensen, A. P. (2020). Towards a network psychometrics approach to assessment: Simulations for redundancy, dimensionality, and loadings (Unpublished doctoral dissertation). University of North Carolina at Greensboro, Greensboro, NC, USA.

Christensen, A. P., & Golino, H. (2020). Statistical equivalency of factor and network loadings. *PsyArXix*. https://doi.org/10.31234/osf.io/xakez

Christensen, A. P., Kenett, Y. N., Aste, T., Silvia, P. J., & Kwapil, T. R. (2018). Network structure of the wisconsin schizotypy scales–short forms: Examining psychometric network filtering approaches. *Behavior Research Methods*, *50*(6), 2531–2550. https://doi.org/doi:%2010.3758/s13428-018-1032-9

Comrey, A. L., & Lee, H. B. (2016). *A first course in factor analysis*. New York: Routledge.

Cramer, A., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–150. https://doi.org/10.1017/S0140525X09991567

Dalege, J., Borsboom, D., Harreveld, F., Waldorp, L. J., & Maas, H. L. (2017). Network structure explains the impact of attitudes on voting decisions. *Scientific Reports*, *7*(1), 4909.

Deboeck, P. R., Montpetit, M. A., Bergeman, C., & Boker, S. M. (2009). Using derivative estimates to describe intraindividual variability at multiple time scales. *Psychological Methods*, *14*(4), 367–386. https://doi.org/http://dx.doi.org/10.1037/a0016622

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numerique*, *17*(1), 61–84.

Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*, *76*(376), 774–781.

Epskamp, M., S. (2018). Network psychometrics. In B. Irwing Paul (Ed.), *The wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 953–986). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118489772.ch30

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212.

Epskamp, S., & Fried, E. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network pschometrics: Combining network and latent variable models. *Psychometrika*, *82*(4), 904–927.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in r. *Journal of Statistical Software*, *25*(5), 1–54. Retrieved from http://www.jstatsoft.org/v25/i05/

Fenton, N. (2016). The internet of radical politics and social change. In *Misunderstanding the internet* (pp. 173–202). Routledge.

Foygel, R., & Drton, M. (2010). Extended bayesian information criteria for gaussian graphical models. In *Proceedings of the 23rd international conference on neural information processing systems - volume 1* (Vol. 1, pp. 604–612). Vancouver, Canada.

Fried, E., Borkulo, C. D. van, Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441.

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2013). A new look at horn's parallel analysis with ordinal variables. *Psychological Methods*, *18*(4), 454–74. https://doi.org/10.1037/a0030005

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A monte carlo evaluation of weighted community detection algorithms. *Frontiers in Neuroinformatics*, *10*, 45. https://doi.org/10.3389/fninf.2016.00045

Ghanem, B., Buscaldi, D., & Rosso, P. (2019). TexTrolls: Identifying russian trolls on twitter from a textual perspective. *arXiv*, (1910.01340). Retrieved from https://arxiv.org/abs/1910.01340

Glynn, C., Tokdar, S. T., Howard, B., & Banks, D. L. (2019). Bayesian analysis of dynamic linear topic models. *Bayesian Analysis*, *14*(1), 53–80.

Golino, H., & Christensen, A. P. (2019). *EGAnet: Exploratory graph analysis: A framework for estimating the number of dimensions in multivariate data using network psychometrics*. Retrieved from https://CRAN.R-project.org/package=EGAnet

Golino, H. F., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using exploratory graph analysis. *Intelligence*, *62*, 54–70.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS One*, *12*(6), e0174035.

Golino, H., Shi, D., Garrido, L. E., Christensen, A. P., Nieto, M. D., Sadana, R., … Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial.

Psychological Methods. https://doi.org/https://doi.org/10.1037/met0000255

Guttman, L. (1953). Image theory for the structure of quantitative variates. *Psychometrika*, *18*(4), 277–296.

Hallquist, M. N., Wright, A. G., & Molenaar, P. C. (2019). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research*, 1–25. https://doi.org/10.1080/00273171.2019.1640103

Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, K., Martinez-Hernandez, V., Perez-Meana, H., Olivares-Mercado, J., & Sanchez, V. (2018). Social sentiment sensor in twitter for predicting cyber-attacks using l1 regularization. *Sensors*, *18*(5), 1380.

Horibe, Y. (1985). Entropy and correlation. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 641–642.

Hornik, K., & Grun, B. (2011). Topicmodels: An r package for fitting topic models. *Journal of Statistical Software*, *40*(13), 1–30.

Hou-Liu, J. (2018). Benchmarking and improving recovery of number of topics in latent dirichlet allocation models. *viXra*. Retrieved from https://vixra.org/abs/1801.0045

Kjellstrom, S., & Golino, H. (2019). Mining concepts of health responsibility using text mining and exploratory graph analysis. *Scandinavian Journal of Occupational Therapy*, *26*(6), 395–410.

Lauritzen, S. L. (1996). *Graphical models* (Vol. 17). Oxford: Clarendon Press.

Libicki, M. C. (1995). *What is information warfare?* The Center for Advanced Command Concepts; Technology, National Defense University. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a367662.pdf

Linvill, D. L., Boatwright, B. C., Grant, W. J., & Warren, P. L. (2019). "THE russians are hacking my brain!" Investigating russia's internet research agency twitter tactics during the 2016 united states presidential campaign. *Computers in Human Behavior*, (99), 292–300.

Linvill, D. L., & Warren, P. L. (2018). *Troll factories: The internet research agency and state-sponsored agenda building*. Clemson University. Retrieved from https://pwarren.people.clemson.edu/Linvill_Warren_TrollFactory.pdf

Llewellyn, C., Cram, L., Favero, A., & Hill, R. L. (2018). Russian troll hunting in a brexit twitter archive. In *Proceedings of the 18th acm/ieee on joint conference on digital libraries* (pp. 361–362).

Massara, G. P., Di Matteo, T., & Aste, T. (2016). Network filtering for big data: Triangulated maximally filtered graph. *Journal of Complex Networks*, *5*(2), 161–178.

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. M. (2002). Dynamic factor analysis models for representing process in multivariate time-series. In D. S. Moskowitz & S. L. Hershberger (Eds.), *Multivariate applications book series. Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 235–265). Lawrence Erlbaum Associates Publishers.

Nikita, M. (2016). *Ldatuning: Tuning of the latent dirichlet allocation models parameters* (R package version 1.0.0). https://doi.org/https://CRAN.%20R-project.%20org/package=%20ldatuning

Nikita, M. (2019). *Ldatuning: Tuning of the latent dirichlet allocation models parameters*. Retrieved from https://CRAN.R-project.org/package=ldatuning

Phan, X.-H., Nguyen, L.-M., & Horiguchi, S. (2008). Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In *Proceedings of the 17th international conference on world wide web* (pp. 91–100).

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *J. Graph Algorithms Appl.*, *10*(2), 191–218.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130–137.

Rajadesingan, A., & Liu, H. (2014). Identifying users with opposing opinions in twitter debates. In *International conference on social computing, behavioral-cultural modeling, and prediction* (pp. 153–160). Springer.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Roeder, O. (2018). Why we're sharing 3 million russian troll tweets. *FiveThirtyEight*, *Retrieved from https://fivethirtyeight.com/features/why-were-sharing-3-million-russian-troll-tweets/*. Retrieved from https://fivethirtyeight.com/features/why-were-sharing-3-million-russian-troll-tweets/

Rosenstein, M. T., Collins, J. J., & De Luca, C. J. (1993). A practical method for calculating largest lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, *65*(1-2), 117–134. https://doi.org/10.1016/0167-2789(93)90009-P

Singh, N., Hu, C., & Roehl, W. S. (2007). Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development. *International Journal of Hospitality Management*, *26*(1), 131–147.

Stewart, L. G., Arif, A., & Starbird, K. (2018). Examining trolls and polarization with a retweet network. In *Proc. ACM wsdm, workshop on misinformation and misbehavior mining on the web*.

Szafranski, R. (1995). *A theory of information warfare: Preparing for 2020*. Air University Maxwell Airforce Base. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a328193.pdf

Taddeo, M. (2017). Cyber conflicts and political power in information societies. *Minds and Machines*, *27*(2), 265–268.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Lecture notes in mathematics (vol. 898)* (pp. 366–381). Springer. https://doi.org/10.1007/BFb0091924

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Van Der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861.

Van Der Maas, H. L., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, *5*(2), 16. https://doi.org/10.3390/jintelligence5020016

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327.

Whitney, H. (1936). Differentiable manifolds. *The Annals of Mathematics*, *37*(3), 645–680. https://doi.org/10.2307/1968482

Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*(3), 263–311.

Williams, D. R., & Rast, P. (2019). Back to the basics: Rethinking partial correlation network methodology. *British Journal of Mathematical and Statistical Psychology*, 1–25. https://doi.org/doi:%2010.1111/bmsp.12173

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, *30*(5), 316–327.

Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). Disinformation warfare: Understanding state-sponsored trolls on twitter and their influence on the web. In *Companion proceedings of the 2019 world wide web conference* (pp. 218–226).

Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). Who let the trolls out? Towards understanding state-sponsored trolls. In *Proceedings of the 10th acm conference on web science* (pp. 353–362).

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 377–402. https://doi.org/doi.org/10.1080/10705510802154281

Ziegler, C. E. (2018). International dimensions of electoral processes: Russia, the usa, and the 2016 elections. *International Politics*, *55*(5), 557–574.