Concept Drift detection on Social Network Data using Cross-Recurrence Quantification Analysis

Rodrigo F. de Mello,[1, a)] Ricardo A. Rios,[2, b)] Paulo A. Pagliosa,[3, c)] and Caio S. Lopes[1, d)]
[1)] *Institute of Mathematics and Computer Science, University of São Paulo, São Carlos, SP, Brazil, 13566-590.*
[2)] *Department of Computer Science, Federal University of Bahia, Salvador, BA, Brazil, 40170-110.*
[3)] *FACOM, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil, 79070-900.*

(Dated: 22 May 2018)

Motivated by the Brazilian demonstrations against the federal government in 2013, we designed and started the development of the TSViz (Time Series Visualization) project, which aims at summarizing the public opinion through the analysis of short messages posted in Twitter. The first version of TSViz was made available on december 2017, presenting the results of text processing in terms of novelty detection, sentiment analysis, most relevant words, geopositioning of Twitter's users, etc. Such version also includes a Concept Drift Detection tool to point out the most relevant moments of behavior change for time series, which is based on Cross-Recurrence Quantification Analysis as detailed in this article.

Keywords: Cross-Recurrence Quantification Analysis, Nonlinear Time Series, Social Networks.

The last decades have been characterized by the design and development of new technological tools to allow faster communication among people. In that context, the social network Twitter, created in 2006, reached an enormous popularity around the World as a microblogging service, where a user under a given nickname can publish short messages called tweets. In 2015, we started the development of the TSViz project to model the public opinion (or ideological biases) over time and how they relate to different topics. TSViz is composed of a module that continuously collect tweets under hashtags of interest, while another one is responsible for processing texts and transform them into different time series. TSViz time series map different sorts of analysis, including, but not limited to: novelty detection, sentiment analysis, most relevant words, geopositioning, etc. While analyzing time series, we decided to immerse them into phase spaces, using Takens' embedding theorem, and the Cross-Recurrence Quantification Analysis to compare how trajectories evolve by assessing the longest diagonal line. From that, we obtained a Concept Drift detection approach to point out the most relevant events associated with a given hashtag. For example, while analyzing the novelty level each new tweet adds to the historical information for hashtag dilmabr (the one used by the Brazil's former president), we observed the most relevant events were associated to the exact date the impeachment was voted in the House of Representatives, the exact date Petrobras (the main Brazilian oil company) confirmed the corruption to the stockholders, and the date Fidel Castro passed away (one of her main allies). We have now been constantly working on improving the TSViz web interface (`http://www.tsviz.com.br`) to provide such features to end users.

I. INTRODUCTION

The last decades have been characterized by the design and development of new technological tools to allow faster communication among people around the World. In that context, Twitter, a social network tool created in 2006, has a prominent success in providing short communications from one person to his/her followers. Currently, Twitter has more than 320 million of monthly-active users (`http://twitter.com/`). Users' activities are referred to as tweets and represent the publication of short texts (at most 140 characters), pictures, videos, links, and retweets (repost of someone else's tweet).

Tweets are written by users under a given nickname who may also add hashtags to their messages. Twitter provides an Application Programming Interface (API) so one can write a client-side application to monitor specific hashtags. By monitoring those subjects, one can attempt to model the public opinion (or ideological biases) along time and how they relate to different topics. For example, one could monitor the political or economical situation: i) in Brazil to provide evidence about how people have been facing particular issues; ii) in Syria to understand the war impacts on people and how other factors, such as health and refugees supports has been provided.

In this context, we propose TSViz[15,16] (available at

[a)]Electronic mail: mello@icmc.usp.br
[b)]Electronic mail: ricardoar@ufba.br
[c)]Electronic mail: pagliosa@facom.ufms.br
[d)]Electronic mail: caiodesalopes@gmail.com

http://www.tsviz.com.br), a Web tool that monitors tweets and builds up data streams to be later processed. TSViz was designed to plug new components which are responsible for generating time series. For example, consider tweets associated to hashtags '#Brazil' and '#impeachment' have been collected along one year. TSViz is capable of processing a sequence of tweets under a given subject to produce: i) time series with information novelties, helping us to understand how tweet content changes along time; ii) time series with the frequencies of tweets per day; iii) time series with the positive or negative sentiments that users stated along their messages; iv) geopositioning system of users according to their profiles; v) time series with the novelty degree of most common words used throughout messages. This list of features is not extensive but simply an example of the components provided by TSViz.

In addition to build up those series, TSViz also provides data analysis tools so that users are capable of understanding tendencies and relationships among different subjects. Moreover, TSViz offers a Web tool so users will visualize data evolution. From this perspective, we intend to support researchers from different areas of knowledge to understand how real-world situations are mapped or impact in the Social Media and in the public opinion.

We are not interested in discussing all contributions of the TSViz project in the context of this paper, but only (iii) listed next: i) the visualization of data streams/time series; ii) the use of Dynamical Systems, Statistics and Machine Learning tools to provide relevant information about social tendencies, with particular interest of our group on hashtags related to politics, economics and health issues; and iii) the cross-referential analysis among several time series in order to understand how past events affect future ones.

The main contribution of this manuscript comes after evolving the TSViz project in several aspects, which motivated us to employ the Cross-Recurrence Quantification Analysis (CRQA) to compare the trajectories of consecutive data windows in order to point out modifications along time, a.k.a. Concept Drift Detection in the context of Data Streams[4,6] and Machine Learning[2,14]. At first, we take any of the time series produced by TSViz and set a given window length, which is used to slide along observations. Every data window is then provided as input to CRQA, which employs Takens' immersion theorem[19] to reconstruct such time series into the phase space[1]. A consecutive window will, therefore, result in a further phase space to be compare against the previous one by using the CRQA measurements[5,20,21]. In particular, in this paper, we employ the maximum diagonal line to point out similarities between consecutive data windows and, consequently, detect when a significant data drift happens[17].

As main contribution, we decided to analyze the most recent data collected by TSViz in order to detect relevant drifts, related to hashtags from politics as well as others producing collateral impacts in Brazil. Results confirm the drifts detected by our approach are directly related to the main fact published by TV channels, newspapers and the Web in general.

This paper is organized as follows: Section II introduces the TSViz project; Section III provides the necessary background on the Cross Recurrence Quantification Analysis; Results are presented and discussed in Section V; Section VI draws conclusions and, finally, references are listed.

II. THE TSVIZ PROJECT

Twitter's users daily produce a huge amount of data which motivated several researchers and companies to design tools and algorithms to collect and analyze tweets, aiming at extracting patterns to understand users' behavior and feelings about specific subjects or trends[3,7,8,10,11,22]. However, most of such analysis focused on creating a dataset of collected tweets and analyzing them by using offline approaches. This is the main gap that motivated us to propose the TSViz (Time Series Visualization) project, a Web tool designed to analyze information gathered from Twitter.

Figure 1 illustrates the main components involved in the TSViz architecture: i) TSViz Robot; ii) TSViz Core; and iii) TSViz User Interface. TSViz Robot online collects every tweet published by users according to monitored hashtags or usernames.

We are already in the second version of this robot designed using the Twitter API (http://developer.twitter.com/en.html), which provides read-and-write access from/to Twitter servers. Using this API, we developed a listener that registers, in the Twitter servers, a set of hashtags and users to be monitored (Figure 1 – Arrow 1). Hence, once any tweet around the World is published using the monitored hashtags and/or users, the robot receives a message containing all information about them (Figure 1 – Arrow 2). In this manner, every hashtag or username defines a new multidimensional data stream/time series. Finally, TSViz Robot stores the tweet information into our databases to be later analyzed by the TSViz Core (Figure 1 – Arrow 3).

After monitoring and collecting tweets, the robot saves them as a multidimensional data stream/time series in our databases to be later analyzed. This data stream contains all tweets information, including: creation timestamp, owner, greographic location, language, the tweet itself and all other parameters provided by Twitter. Then another TSViz component, called TSViz Core, is responsible for performing two main tasks: i) firstly, a producer module periodically looks for and incrementally converts all tweets associated either to a hashtag or a username into time series (Figure 1 – Arrow 4). Then, the resultant time series is also stored into our databases using different time resolutions through a Discrete Wavelet Transform (Figure 1 – Arrow 5), so they can be later analyzed by

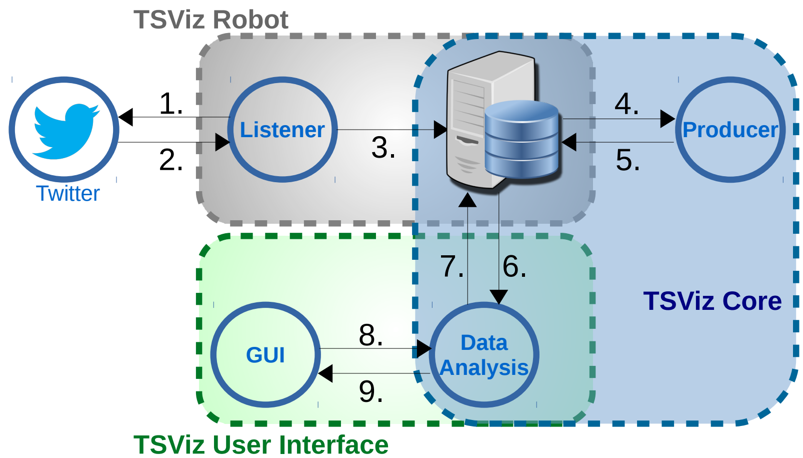


FIG. 1: TSViz Architecture.

the producer module.

Currently, we have the following producer modules: i) a novelty detection module based on the Normalized Compression Distance (NCD)[12], to inform about binary changes along messages under the same context; ii) a cumulative module that sums up series observations along time to discover data tendencies. This can be employed in conjunction with any of the other components; iii) a module to extract the most relevant words in a given context along time; iv) a module to perform the correlation among different subjects using the set of their most tweeted words along time; v) a novelty detection module, based on Shannon's entropy[18], to point out changes in the set of words under a subject; v) a geopositioning module that takes users' profiles and associate them with the most probable place of tweeting; vi) a sentiment analysis module to measure the level of positiveness of negativeness of posted messages; and, finally, vii) a concept drift detection module based on the Cross-Recurrence Quantification Analysis (CRQA)[5,13,21] that informs the most important moments data changes along time, allowing to point out modifications in the generation processes[9]. The next section details CRQA, and afterwards our approach to detect drifts in time series from social networks is introduced.

III. CROSS-RECURRENCE QUANTIFICATION ANALYSIS

Recurrence analysis tools are useful to study and characterize the behavior of dynamical systems by reconstructing the produced data in phase spaces[13]. The reconstruction of such spaces is typically performed using the Takens' immersion theorem[19] that allows to analyze system states from which one may obtain a regression function to model and forecast future states[5]. Among those tools, Recurrence Plot (RP) support the study and analysis of time series recurrences[5,13,21]. In summary, RP produces an output matrix, referred to as Recurrence Matrix, representing how close two states are in phase space, thus allowing to understand how states evolve along time. Equation 1 is employed to measure nearby states, in which ε is a distance threshold, $||\cdot||$ corresponds to the norm employed to calculate distances among states i and j, and, finally, $\Theta(\cdot)$ is a heaviside function (Equation 2).

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - ||\vec{x_i} - \vec{x_j}||), \{i, j = 1, 2, \ldots, N\} \quad (1)$$

$$\Theta(\alpha) = \begin{cases} 0, & \alpha < 0 \\ 1, & \alpha \geq 0 \end{cases} \quad (2)$$

To proceed with the RP analysis, the recurrence matrix is plotted by using black dots when $R_{i,j} = 1$ and white ones when $R_{i,j} = 0$. Structures present in RP provide important information on the dynamical systems under study. For instance, isolated points mean system states are rarely repeated, emphasizing a stochastic behavior. On the other hand, diagonal lines occur when there is persistent behavior, showing system states are highly recurrent.

As previously discussed, the RP analysis is performed on a single time series unfolded on its phase space. However, there exists scenarios in which it is important to analyze the relationship between states from two different systems. Aiming at overcoming this situation, Cross Recurrence Plots (CRPs) are a natural complement of RPs[21], which intend to study the dynamic of two systems by representing them into the same phase space, looking for occurrences when state of a system recurs to state of the other one[13]. Similarly to RP, CRP also produces a matrix, whose recurrence is defined by Equation 3, in which $\vec{x}$ and $\vec{y}$ represent the trajectories of two dynamical systems in a d-dimensional phase space[13,20].

$$CR_{i,j}^{\vec{x},\vec{y}}(\varepsilon) = \Theta(\varepsilon - ||\vec{x_i} - \vec{y_j}||), \{i = 1, 2, \ldots, N\}, \{j = 1, 2, \ldots, M\} \quad (3)$$

The structures produced by CRP are analyzed by the Cross Recurrence Quantification Analysis (CRQA), which provides a set of measures to quantify, for example, the determinism rate, the recurrence rate, the maximal diagonal line, etc. In this article, we are particularly interested in the maximal diagonal line ($L_{\max}$) measure, whose results are shown in Section V. The length of a diagonal line (l) is related to the time steps in which two trajectories in phase space are close to each other, emphasizing similar behavior. By measuring the maximal diagonal line in a RP plot, we can analyze the exponential divergence of the phase space trajectory. In summary, the faster the trajectory segments diverge, the shorter are the diagonal lines[13,20]. The maximal diagonal line ($L_{\max}$) is calculated by Equation 4, in which $P(l)$ is the frequency distribution of diagonal lines with length equal to l.

$$L_{\max} = max(\{l_i\}_{i=1}^{N_l}) \quad (4)$$

$$N_l = \sum_{l \leq l_{min}} P(l) \quad (5)$$

In the next section, we present the approach in which this CRQA measure was applied to compute the similarity among texts published in Twitter and to analyze how the public opinion on specific topics has changed over time.

IV. EXPERIMENTAL SETUP

As previously discussed in Section I, we used CRQA to understand how people reactions in Twitter evolves along time and how that is connected to important facts from the real world. In this sense, we selected a set of Twitter hashtags and users to be monitored by TSViz. It is important to highlight such hashtags and users are not related to the authors' political view, but to the main terms used by the general media. The main motivation to select them was to understand how they are connected to the severe issues faced in Brazil, leading to a strong and highly frequent reaction on Twitter.

As these hashtags and users have been monitored by TSViz Robot, the collected texts (for any upper or lower written form) are transformed into time series by the TSViz Core component. Although there exists several producers, in this manuscript we focused on the Normalized Compression Distance (NCD) and Sentiment Analysis (SA).

In summary, NCD allows to compute the similarity among texts published in Twitter. By using this producer, we can measure if new textual content has been shared or basically the same (as retweeting). The resultant time series is obtained by applying NCD on pairs of consecutive tweets as defined in Equation (6).

TABLE I: The set of hashtags and users analyzed by TSViz.

Hashtag/User	Short description
dilmabr	Twitter account of the former Brazil's president Dilma Rousseff
LulaPeloBrasil	Twitter account of Luis Inácio Lula da Silva, also a former president
MichelTemer	Twitter account of the current president of Brazil Michel Temer
ForaTemer	Hashtag against Michel Temer
jairbolsonaro	Twitter account of a controversial figure in Brazil's political scenario
petrobras	a semi-public Brazilian multinational corporation in the petroleum industry
SenadoFederal	Federal Upper House of the National Congress of Brazil
CamaraDeputados	Lower Chamber of Deputies
politica	Hashtag associated with politics in Brazil
STF_oficial	Twitter account of the Supreme Federal Court of Brazil
dolar	Portuguese term for "dollar"
dengue	Term associated with the disease "dengue"
Zika	Term associated with the disease "zika"
saude	Portuguese term for "health"

$$X_N = \{\mathrm{NCD}(t_1, t_2), \ldots, \mathrm{NCD}(t_i, t_{i+1}), \ldots, \mathrm{NCD}(t_{T-1}, t_T)\}. \quad (6)$$

In this equation, T is the total number of collected tweets at the current time instant. The $NCD(\cdot)$ function is presented in Equation 7, such that $Z()$ is the binary length of a compressed file in bytes. The compression tool adopted in our experiment was BZip2. In our case, a file contains either the text of a single tweet ($Z(t_i)$) or concatenated texts of two consecutive tweets ($Z(t_i, t_j)$).

$$\mathrm{NCD}(t_i, t_j) = \frac{Z(t_i t_j) - \min\{Z(t_i), Z(t_j)\}}{\max\{Z(t_i), Z(t_j)\}}. \quad (7)$$

Our Sentiment Analysis module was implemented using a Naïve Bayes classifier[2,14]. We implemented such classifier to employ a training stage, in which the probabilities of words published along tweets are used to label unknown posts as being negative or positive in a continuous range $[-1, 1]$. Then, we only classify tweets daily published, calculating their probability of positiveness and negativeness by using Equations 8 and 9, respectively. They compute the product of probabilities that every word w has for wither supervised label (positive or negative). As a consequence, we obtain a time series summarizing positive and negative feelings about a given hashtag or username along time using Equation 12, in which $\mathrm{SA}(t_i) \in [-1, 1]$ corresponds to the overall sentiment analysis for a single tweet. Equations 10 and 11 are used to compute such time series, in which we basically normalize the positive and negative probabilities to sum up to one.

$$P(t_i, \text{Positive}) = \prod_{w \in t_i} P(w, \text{Positive}). \tag{8}$$

$$P(t_i, \text{Negative}) = \prod_{w \in t_i} P(w, \text{Negative}). \tag{9}$$

$$\text{NP}(t_i) = \frac{P(t_i, \text{Positive})}{P(t_i, \text{Positive}) + P(t_i, \text{Negative})}. \tag{10}$$

$$\text{NN}(t_i) = \frac{P(t_i, \text{Negative})}{P(t_i, \text{Positive}) + P(t_i, \text{Negative})}. \tag{11}$$

$$\text{SA}(t_i) = \text{NP}(t_i) - \text{NN}(t_i) \tag{12}$$

After obtaining the time series generated by NCD and Sentiment Analysis for every tweet in our list of hashtag/user (see Table I), we designed an approach that uses CRQA to detect concept drifts. In summary, this approach runs two straightforward steps. Firstly, we define a initial time window ω_i containing a set of observations from the time series produced by either NCD or Sentiment Analysis. Then, the window is slid considering a time interval $\omega_{i+\tau}$ and a new set of observation is extracted. Secondly, the two windows are analyzed by using CRQA ($CR_{i,j}^{\overrightarrow{\omega_i},\overrightarrow{\omega_{i+\tau}}}(\varepsilon)$) and the recurrence rate is stored into a new time series. The process is, finally, repeated by sliding the windows up to reach the last observation. The new time series will keep the similarity between pairs of windows and, if a difference is noticed, then a concept drift detected, i.e., a new information was detected or the people's sentiment has changed. The results obtained with this new approach are presented and discussed in the next section.

V. RESULTS

We start with the analysis of the former Brazil's president, Dilma Rousseff. According to CRQA, whose maximal diagonal lines $L_{\max}$ along the data windows are provided in Figure 2, we notice very relevant events on November 4th and 5th, 2017 for the NCD series (this means CRQA was computed on NCD and $L_{\max}$ was plotted to illustrate our conclusions). That was motivated by her tweets in which she declared male prejudice against women during her presidential term. On the same period, one of her followers posted a message claiming she is criminal, reinforcing novelties and, consequently, the drift detection found.

Luiz Inácio Lula da Silva, a.k.a. Lula, was president of Brazil from 2003 to 2011. He is a founding member of the Workers' Party (Partido dos Trabalhadores – PT) and referred to as one of the most popular politicians

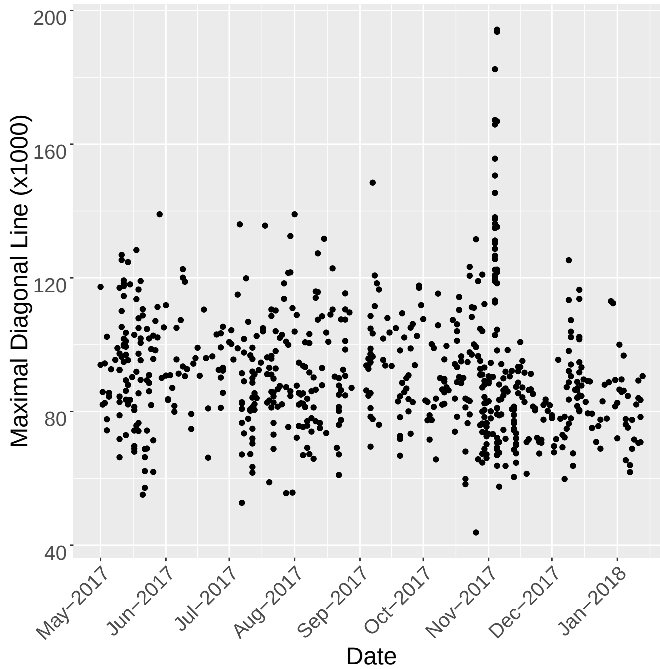


FIG. 2: Maximal Diagonal Line for "dilmabr" computed using the Normalized Compression Distance.

in the history of Brazil. Lula was convicted of money laundering and passive corruption, defined in Brazilian criminal law as the receipt of a bribe by a civil servant or government official. He was sentenced to nine years and six months in prison by judge Sérgio Moro under the Operation Car Wash (*Operação Lava Jato* in Portuguese), but remains free pending an appeal of such sentence.

There is a particular date, in Figure 3, that calls our attention while analyzing his hashtag, which is July 1st, 2017, the day that Joesley Batista (one of the owners of JBS – the largest, by sales, meat processing company in the world) stated to have given a bribe of R$ 300 million (around US$ 90 million) to Lula.

We also analyzed publications that directly mentioned the current Brazil's president Michel Temer by using CRQA on NCD (Figure 4a). The first higher divergences happened in May 12th, 2017 when he completed 1 year as president after the impeachment of Dilma Rousseff. A week later, a new peak was detected when the local media released a secret recording in which Michel Temer is discussing hush money with Wesley Batista, a chief executive from one of JBS.

In May 28th, 2017, a few days before the judgment of his impeachment, Michel Temer changed the minister of justice indicating a person that has a good relationship with the Superior Electoral Court and the Supreme Federal Court. People faced this unconventional move, performed by Temer, as a way to defend himself against corruption allegations.

In July 11th, 2017, two important events happened involving Temer. In the first one, Brazil's senate approved

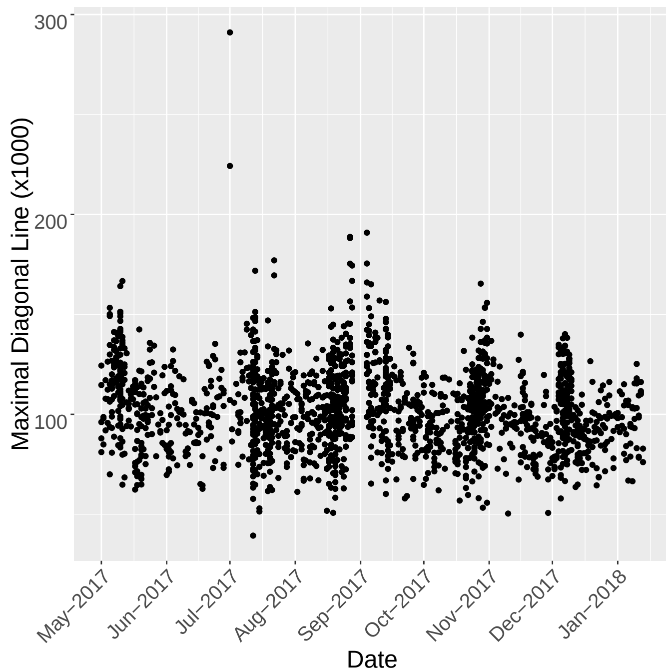


FIG. 3: Maximal Diagonal Line for "LulapeloBrasil" computed using the Normalized Compression Distance.

a significant overhaul of labor rules. The second one happened after the rapporteur of the lower house committee, that was examining a corruption charge against Temer, recommended to vote putting him on impeachment trial. However, in July 15th, 2017, the lower house committee rejected the corruption charge, triggering another divergence spike. In September 2017, NCD emphasized another divergence when Temer visited China and stated the new labor rules had reduced the number of unemployed people. Finally, other spikes were detected in December 2017 when Temer announced his next priority was to approve pension reform. As the labor, the pension reform faced several polemical issues, causing strong reactions on Twitter.

The sentiment analysis performed on tweets mentioning the current Brazil's president Michel Temer has presented similar divergences when compared to NCD. In summary, our concept drift analysis detected spikes related to people reaction when Temer completed 1 year in presidency, also during his judgement in the lower house, and while announcing the pension reform as shown in Figure 4b.

Several detections were produced for the hastag "foratemer" (get out Michel Temer) on the following dates April 5th-6th, July 6th-9th, August 2nd, and September 10th, 2017, while analyzing NCD (Figure 5a). On April 5th-6th, 2017, a worker painted the message "Fora Temer" on the walls of the National Museum in Brasília, the capital of Brazil. On July 6th-9th, 2017, several artists started a campaign in attempt to impeach Temer. On August 2nd, 2017, a group of deputies get into the lower chamber with a band and posters asking for the impeachment. On September 10th, 2017, a famous Brazilian musician, Gilberto Gil, stated Michel Temer had been already impeached after a gig that commemorated 40 years of his classic album "Refavela". The same dates are confirmed while analyzing SA (Figure 5b).

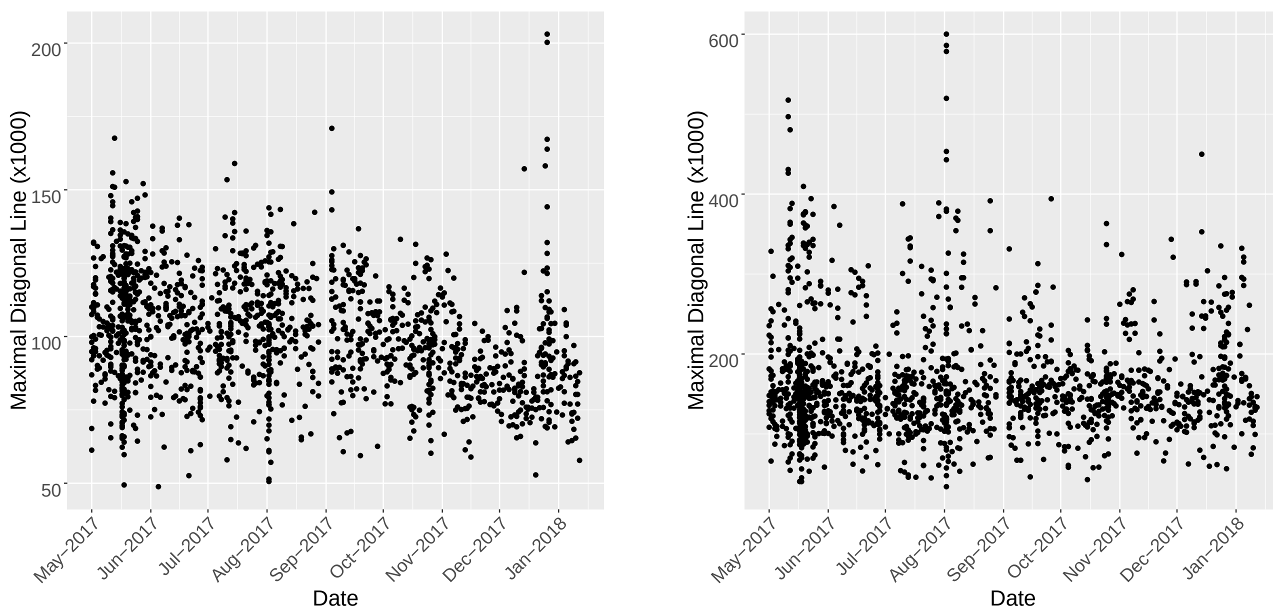
A different behavior is observed while analyzing the detections for Jair Bolsonaro, a Brazilian politician and former military officer who was elected into the Chamber of Deputies. He is seen as a controversial figure in Brazil, due to his positions against the left wing. He has been visiting several cities around Brazil, being acclaimed by some and hated by others. That is in fact observed while analyzing NCD and SA (Figure 6), given no evident and single detection, but instead several changes along time. That perception is the same collected along news and Youtube about this public person.

Petrobras (Petróleo Brasileiro SA) is a semi-public Brazilian multinational corporation in the petroleum industry. In 2014, the largest corruption scandal in the history of Brazil was centered around Petrobras. Here, we analyze the concept drift of NCD (Figure 7) whose main dates are: October 20th-21st, 2017. On those particular dates the operation Car Wash launched ten court orders resulting from the payment of bribes from Odebrecht (Brazilian conglomerate consisting of diversified businesses in the fields of engineering, construction, chemicals and petrochemicals).

Next, we analyzed the Federal Senate, which is the upper house of the National Congress of Brazil. In terms of NCD drifts (Figure 8a), we found the following important dates: July 4th-5th and December 25th, 2017. The motivation for drifts on July 4th-5th, 2017 are due to the senate had voted the labor amendment to change employment contracts in Brazil and to make them move a little more towards a liberal format. On December 25th, 2017 several news were published about involvements of the main politicians from the Senate with the Operation Car Wash. In terms of SA (Figure 8b), the main dates were December 19th-26th, 2017, which are related to the same subjects but also including a public petition for the impeachment of Gilmar Mendes, one of the members of the Supreme Court of Brazil, who has released important politicians from prison.

The Chamber of Deputies composes the federal legislative lower house of the National Congress of Brazil. Such a chamber comprises 513 deputies elected in a proportional fashion to represent each state in a four-year term. Figures 9a and 9b shows the concept drift detection results computed on two respective time series: the Normalized Compression Distance (NCD) and the Sentiment Analysis (SA).

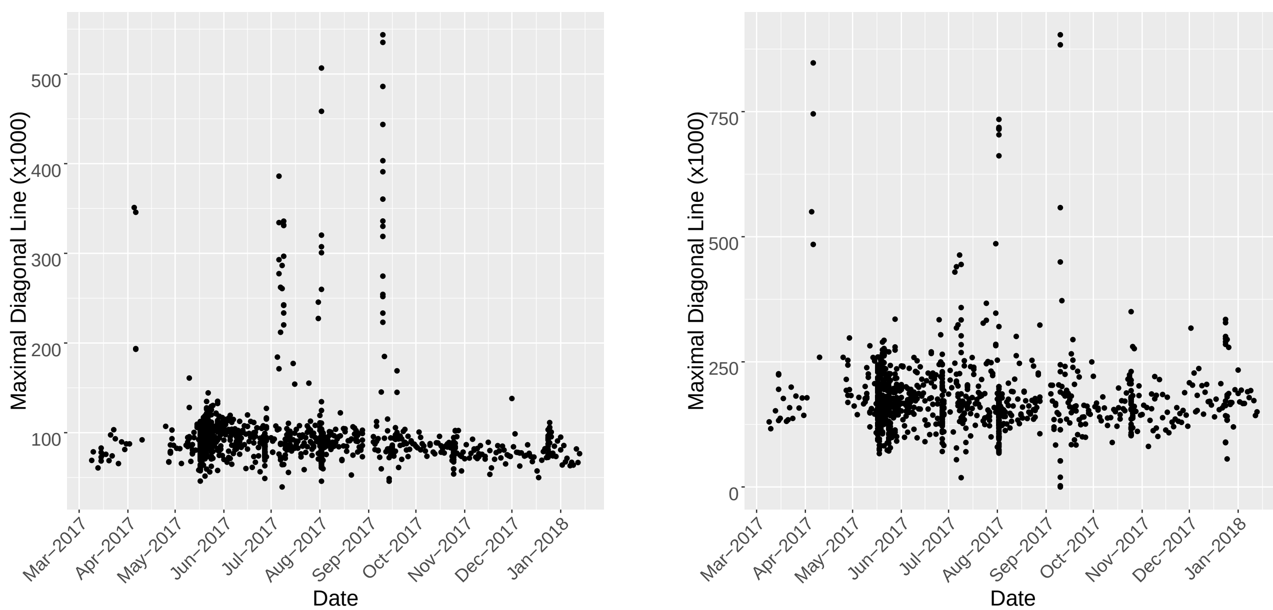
In case of NCD (Figure 9a), the most relevant dates are July 30th, November 2nd and December 19th, 2017. On July 30th, 2017, the Chamber started discussing about the allegations against the current Brazil's president Michel Temer. On November 2nd, 2017, Edson Fachin, one of the members of the Supreme Federal Court



(a) NCD

(b) SA

FIG. 4: Maximal Diagonal Line for "MichelTemer".



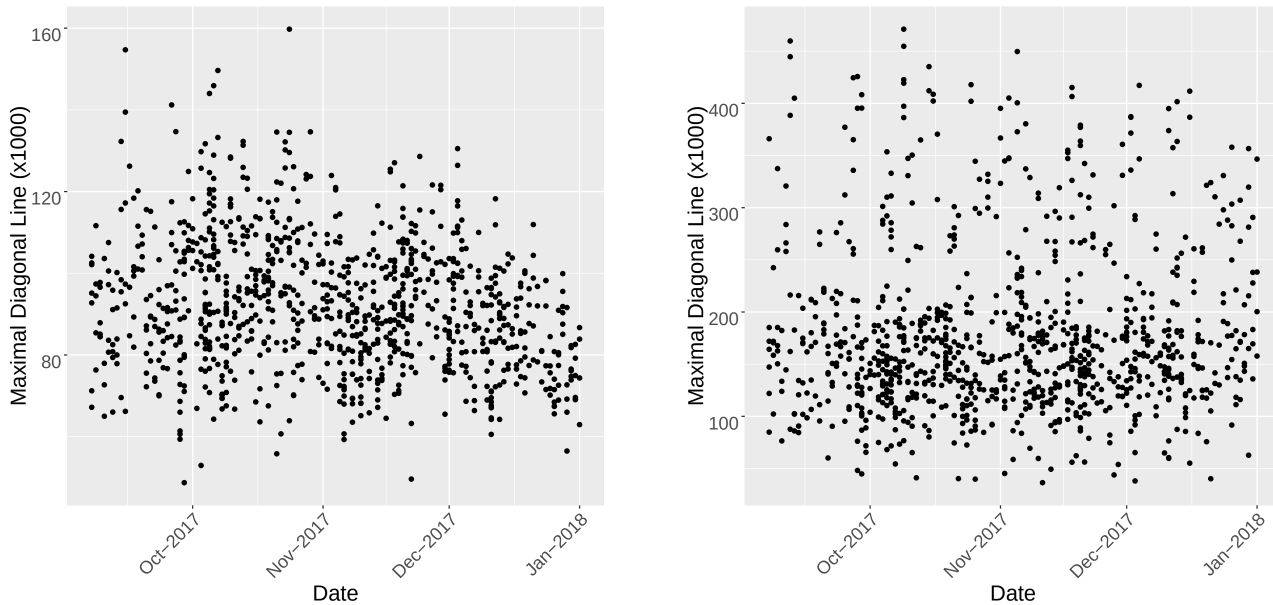
(a) NCD

(b) SA

FIG. 5: Maximal Diagonal Line for "foratemer".

of Brazil made the decision to carry on with a judicial process against Michel Temer and two of his ministers (Eliseu Padilha and Moreira Franco), besides the Chamber of Deputies had denied the authorization to file a suit against the president. On December 19th, 2017, the Chamber approved the lobbying activity in the entities of the federal public administration.

In case of SA (Figure 9b), two of the same dates were also detected: July 30th, November 2nd and December 19th, 2017, confirming the same facts. Although a new date appears as important too: August 25th, 2017, due to the public opinion against corruption, the Chamber of Deputies is pressured to propose new forms of political organizations. On that particular date, they discussed about the public funding of elections, the party coalitions, and barrier clauses to avoid the evergrowing num-



(a) NCD (b) SA

FIG. 6: Maximal Diagonal Line for “jairbolsonaro”.

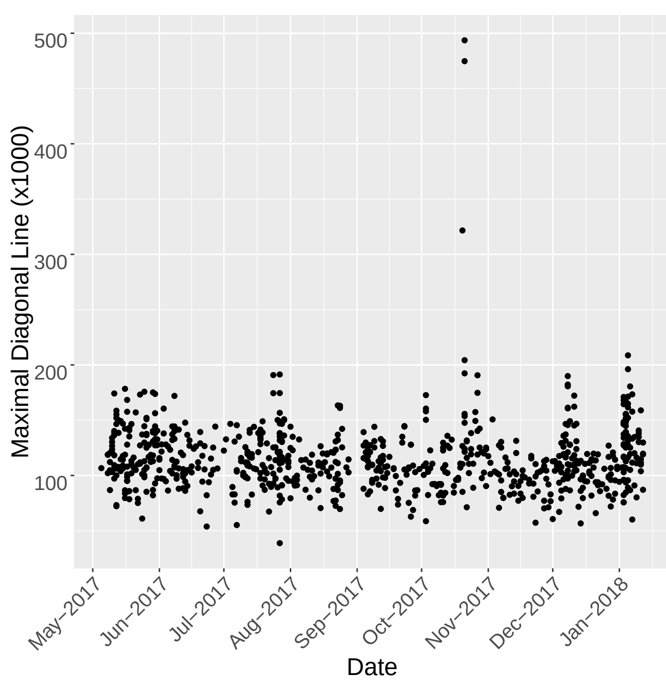


FIG. 7: Maximal Diagonal Line for “petrobras” computed using the Normalized Compression Distance.

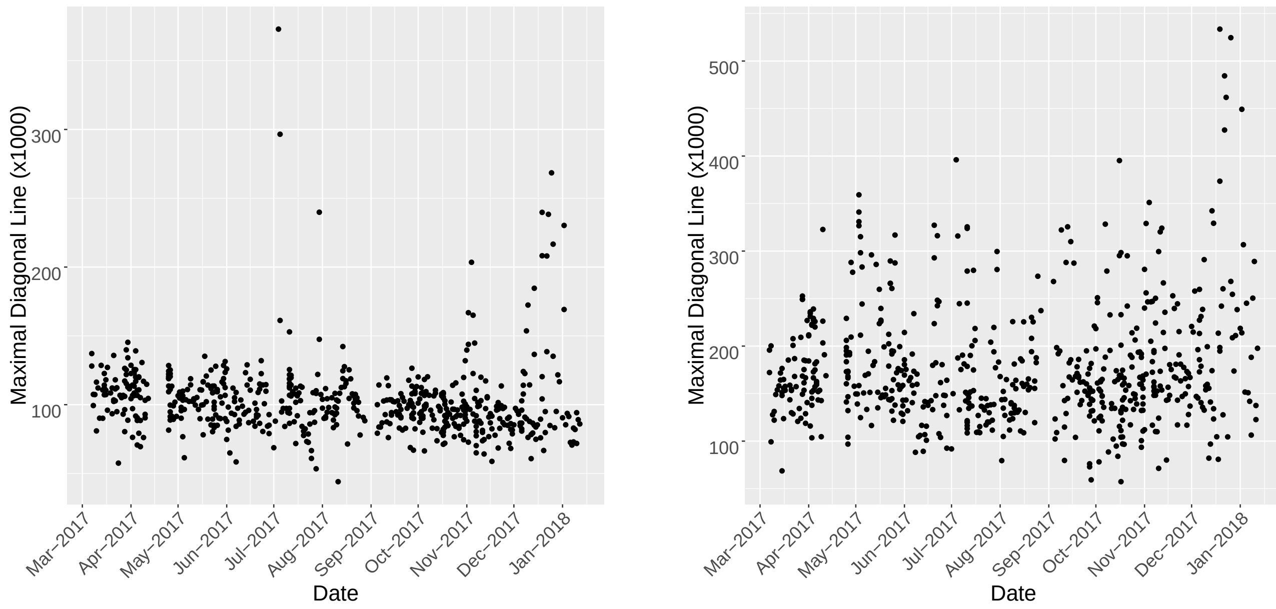
ber of parties.

The Supreme Federal Court of Brazil is the highest court of law in Brazil for constitutional issues and its rulings cannot be appealed. While analyzing NCD and SA (Figure 10), one important date was detected: December 22nd, 2017. On that day, a public petition for the impeachment of Gilmar Mendes, member of such court, was filed with 1.6 million signatures from all around Brazil. It is important to mention that he released some politicians from prison and many Brazilians believe that is due to his interests with other influential companies and politicians.

After analyzing specific users and terms that have been playing an important role in the current political scenario in Brazil, we decided to study the general term “politic” (Figure 11). As a consequence, it was possible to understand relationships among such users and terms, and discover other themes that have been calling the attention of the population in general.

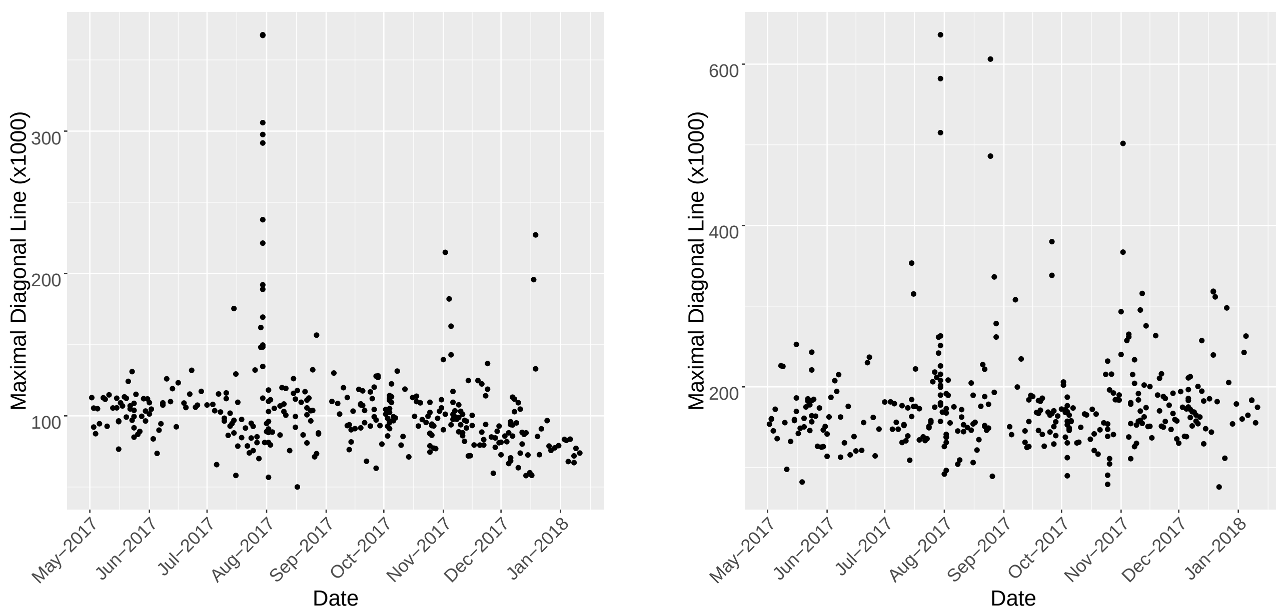
The first observed spike happened in June 4th, 2017 when the local media published analyses on how the political crisis has affected other areas such as education and economy. On August 12th, 2017, a set of news was published pointing out the need of carrying out the political reform, before performing labor and pension reforms. On August 23th, 2017, a NCD divergence emphasized people reactions when the Chamber of Deputies authorized the use of almost 10 billion dollars as budget for the 2018 election, in spite of the severe economical and political crisis in Brazil. On December 2017, new spikes were detected when Temer was submitted a surgical procedure. Finally, a new divergence, not detected by NCD, has showed people reaction when the new Labor Minister of Temer’s government was banned from the position after being found guilty of not paying private workers.

The Brazilian political crisis has directly and strongly affected the local economy. This can be verified by performing a sentiment analysis on tweets mentioning the hashtag “dolar” as shown in Figure 12. A relevant



(a) NCD (b) SA

FIG. 8: Maximal Diagonal Line for “SenadoFederal”.



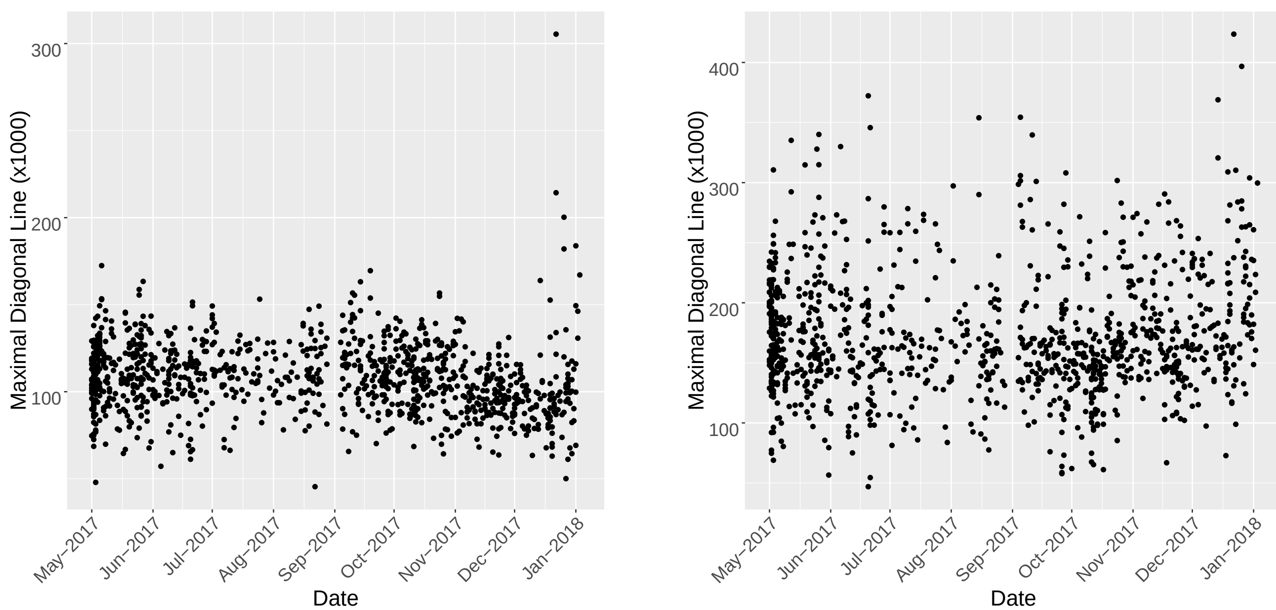
(a) NCD (b) SA

FIG. 9: Maximal Diagonal Line for “CamaraDeputados”.

spike happened in June 19th, 2017 when the current exchange rate between US Dollar (USD) and BRAZIL REAL (BRL) has significantly increased. However, this rate has reduced in August 18th, 2017 as Temer and allies approved some reforms in the congress. On the other hand, in October 16th, 2017, as new corruption cases involving Temer were released, the rate also reacted, increasing its value.

Finally, we also used NCD to analyze tweets published by using the following hashtags “dengue”, “zika” and “saude” (health). Such analyses were useful to understand people’s reaction by facing new symptoms and consequences of diseases transmitted by the aedes aegypti mosquito.

Figure 13 shows the NCD results after analyzing tweets published from May 2017 to January 2017 with the



(a) NCD

(b) SA

FIG. 10: Maximal Diagonal Line for "STF_oficial".

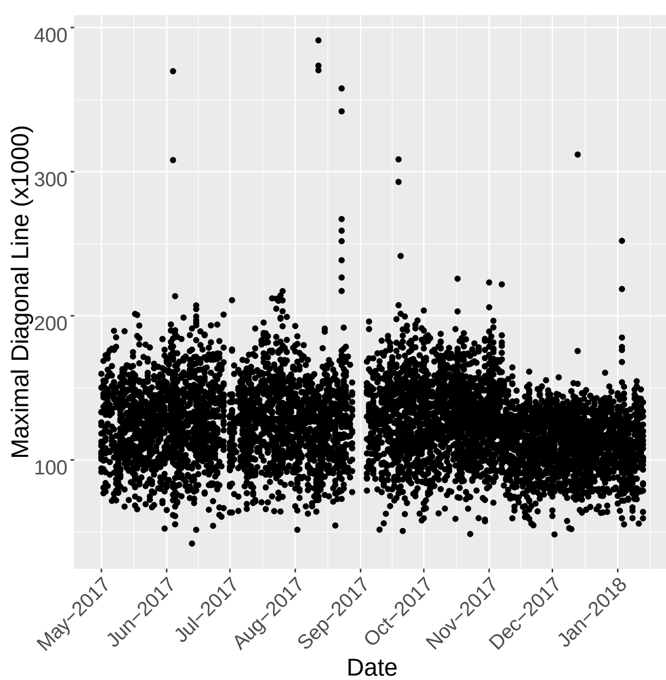


FIG. 11: Maximal Diagonal Line for "politic" computed using the Normalized Compression Distance.

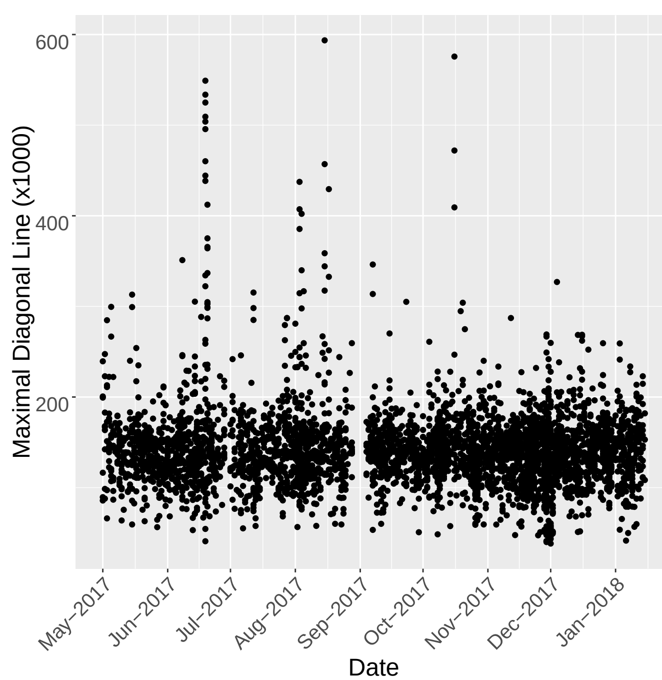


FIG. 12: Maximal Diagonal Line for "dolar" computed using the Sentiment Analysis.

hashtag "dengue". By observing this figure, we correlate the main spikes to important historical facts that happened around the World. The spikes between May 28th-29th, 2017 reflected high epidemic levels in different parts of the World as, for instance, Northeast region of Brazil, Delhi (India), Myanmar, and Sri Lanka. Although dengue fever is a common virus in such regions, an increasing number of the dengue hemorrhagic fever cases, also known as severe dengue, has called the attention of the population. Specially survivors of the devastating floods and landslides of Sri Lanka caused by heavy southwest monsoon during the final week of May 2017.

We also noticed another important spike in August 9th, 2017 when an article published in Nature confirmed the same mosquito species could also transmit the zika virus. On the same day, a campaign was started in Brazil to control dengue as well as to provide public awareness about the habitat and life cycle of the aedes aegypti.

On September 29th, 2017 and November 5th, 2017 new epidemics were registered in Brazil caused by the same mosquito. However, on these dates, the media reported several patients diagnosed with another virus called chikungunya. Finally, new spikes were also detected during November 19th-20th, 2017 when the number of notified cases of dengue has considerably increased in Brazil. This spring period in Brazil is characterized by a greater volume of rain, creating favorable conditions to multiply the mosquito population. For that reason, people reacted in Twitter as the number of reported patients increased, aiming at spreading information on health education during the National Dengue Day.

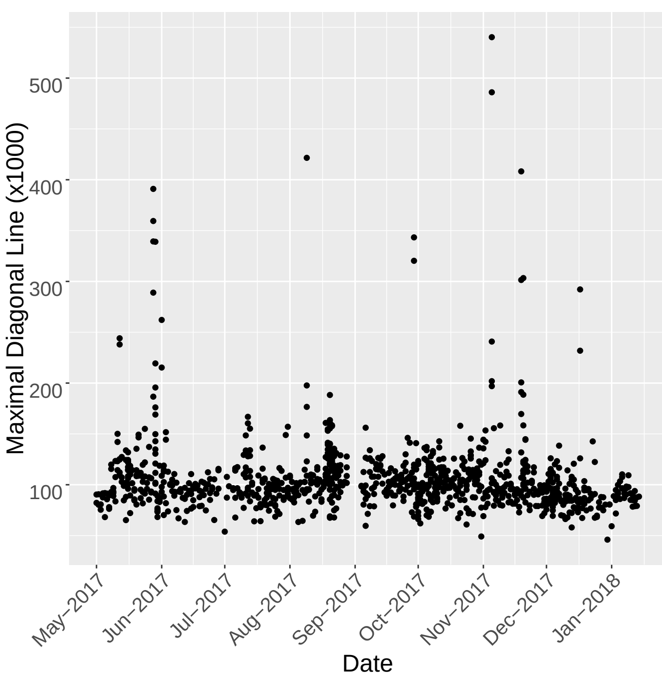


FIG. 13: Maximal Diagonal Line for "dengue" computed using the Normalized Compression Distance.

The zika virus is also transmitted by the same mosquito of the dengue fever. However, this virus is specially aggressive when women are infected in the first trimester of pregnancy, affecting the fetus' central nervous system and leading to malformed babies with microcephaly. By analyzing the NCD results on tweets published using the hashtag "zika" (Figure 14), we noticed four main spikes. The first one happened on July 7th, 2017 when Spanish scientists have discovered a molecule that could be used as a potential drug to cure patients with the zika virus. The second peak summarizes people's reactions on September 17th, 2017 when American researchers published results suggesting the zika virus could kill brain cancer stem cells. On October 10th, 2017, the Brazilian ministry of health showed the uncontrolled mosquito vectors have increased not only the transmission of zika but also malaria, surpassing the previous year by almost 30%. Finally, we noticed a great NCD divergence during October 17th-25th, 2017 when tests to detect the zika virus in blood donners were approved by the US Food and Drug Administration. During the same period, a Brazilian foundation for tropical medicine started a selection to identify patients infected with the zika virus that were able to participate of a study to discover how long the virus could survive in human body.

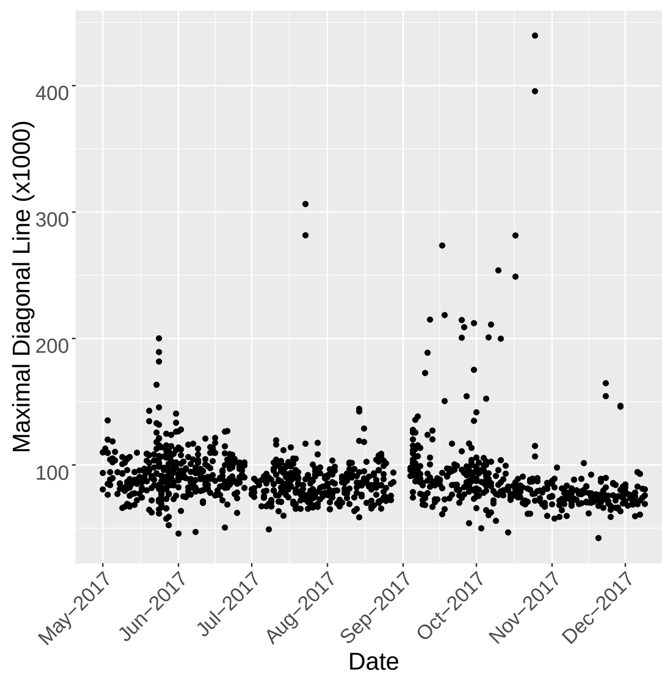


FIG. 14: Maximal Diagonal Line for "Zika" computed using the Normalized Compression Distance.

As last result (Figure 15), we analyzed the NCD divergence computed on tweets published from May 2017 to January 2017, containing a Portuguese word meaning health. We decided to monitor such a word to identify whether the previous hashtags had directly influenced the general publications on health and to understand which are the other concerns related to public health. The first spike was associated to the day on which news reported the public health care system in Brazil has received medicines to treat babies with microcephaly, whose number of cases increased due to the zika virus. The second spike was related to people's reactions by getting to know that new medicines to treat cancer were available in government network of distribution (free of charge). Finally, the last spike occurred during December 22nd-24th, 2017, when the local media published the private health insurance companies had significantly reimbursed the Brazil's public health system (46% greater than in 2016).

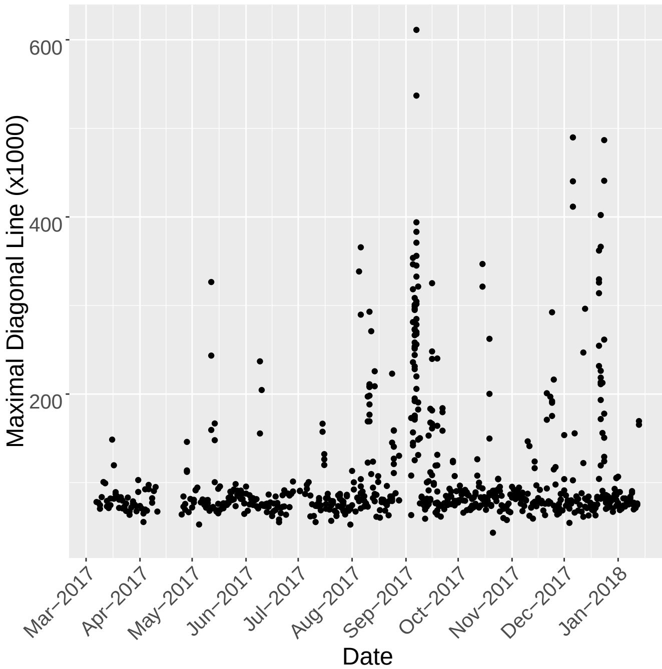


FIG. 15: Maximal Diagonal Line for "saude" computed using the Normalized Compression Distance.

All references supporting our conclusions are listed in the appendix (segmented by hashtag). Any person knowing the Brazil's situation can conclude references are broad enough in terms of political opinions.

VI. CONCLUSION

This manuscript presents an innovative application of Cross-Recurrence Quantification Analysis (CRQA) to understand people's reactions in Twitter while facing real-world issues. In this context, CRQA confirmed to be an important tool to identify implicit information published in social networks. The first information summarizes the similarity rate among different texts. By using NCD (Normalized Compression Distance) and CRQA, we assessed whether most of posts expressed individual opinions or a large sharing of common points of view. According to our results, we noticed different posts are mostly published as new important facts happened in real-world and CRQA helped us to identify such divergences. The second information analyzed in this work was extracted by using a Sentiment Analysis (SA) approach in conjunction with CRQA. After classifying messages according to their positiveness and negativeness, CRQA was used to detect concept drift, in which the public opinion on a specific topic has changed over time. Based on our results, we noticed CRQA and SA allowed to identify important spikes that represented a reaction of Twitter users. Experiments analyzed relevant subjects associated with Brazil, including politics, health, and economics. As future work, we intend to adopt other CRQA measures to extract new information in the context of social network analysis as, for instance, the determinism level to understand how predictable users or hashtags are along time.

ACKNOWLEDGMENT

This paper is based upon projects sponsored by CNPq, Brazil, under grants 441583/2014-8, 303051/2014-0, 302077/2017-0, and FAPESP, Brazil, under grant 2013/07375-0, 2017/16548-6. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of CNPq and FAPESP. The TSViz project has been under registration process according to the Brazilian innovation laws. Anyone interested in using TSViz should contact the authors of this paper.

[1] K. T. Alligood, T. D. Sauer, and J. A. Yorke. *Chaos: An Introduction to Dynamical Systems*. Springer-Verlag, 1997.

[2] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[3] Mondher Bouazizi and Tomoaki Ohtsuki. Opinion mining in twitter how to make use of sarcasm to enhance sentiment analysis. In *IEEE/ACM ASONAM*, pages 1594–1597, 2015.

[4] Lucas de Carvalho Pagliosa and Rodrigo Fernandes de Mello. Applying a kernel function on time-dependent data to provide supervised-learning guarantees. *Expert Systems with Applications*, 71:216 – 229, 2017.

[5] J. P. Eckmann, Oliffson S. Kamphorst, and D. Ruelle. Recurrence plots of dynamical systems. *Europhysics Letters*, 4:973–977, November 1987.

[6] João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. A survey on concept drift adaptation. *ACM Comput. Surv.*, 46(4):44:1–44:37, March 2014.

[7] Wei Gao and Fabrizio Sebastiani. Tweet sentiment: From classification to quantification. In *IEEE/ACM ASONAM*, pages 97–104, 2015.

[8] Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo, and Patrick Meier. TweetCred: A real-time web-based system for assessing credibility of content on twitter. *CoRR*, 2014.

[9] Holgar Kantz and Thomas Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge University Press, Cambridge, 1997.

[10] Shamanth Kumar, Huan Liu, Sameep Mehta, and L. Venkata Subramaniam. Exploring a scalable solution to identifying events in noisy twitter streams. In *IEEE/ACM ASONAM*, pages 496–499, 2015.

[11] Hemank Lamba, Momin M. Malik, and Jürgen Pfeffer. A tempest in a teacup? Analyzing firestorms on twitter. In *IEEE/ACM ASONAM*, pages 17–24, 2015.

[12] Ming Li, Xin Chen, Xin Li, Bin Ma, and P. M. B. Vitanyi. The similarity metric. *Trans. on Info. Theory*, 50(12):3250–3264, 2004.

[13] N. Marwan, M. Romano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6):237–329, January 2007.

[14] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.

[15] Ricardo A. Rios, Caio S. Lopes, Fabio H. G. Sikansi, Paulo A. Pagliosa, and Rodrigo F. de Mello. Analyzing the public opinion on the brazilian political and corruption issues. In *BRACIS 2017*, pages 13–18, 2017.

[16] Ricardo A. Rios, Paulo A. Pagliosa, Renato P. Ishii, and Rodrigo F. de Mello. Tsviz: A data stream architecture to online

collect, analyze, and visualize tweets. In *SAC '17*, pages 1031–1036, New York, NY, USA, 2017. ACM.

[17] Joan Serrà, Xavier Serra, and Ralph G Andrzejak. Cross recurrence quantification for cover song identification. *New Journal of Physics*, 11:093017, Sep. 2009.

[18] C. E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55, January 2001.

[19] Floris Takens. Detecting strange attractors in turbulence Dynamical Systems and Turbulence, Warwick 1980. *Dynamical Systems and Turbulence*, 898:366–381, 1981.

[20] Joseph P. Zbilut, Alessandro Giuliani, and Charles L. Webber. Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, 246(1):122 – 128, 1998.

[21] Joseph P. Zbilut and Charles L. Webber, Jr. Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters A*, 171(3-4):199–203, 1992.

[22] Jun Zou, Faramarz Fekri, and Steven W. McLaughlin. Mining streaming tweets for real-time event credibility prediction in twitter. In *IEEE/ACM ASONAM*, pages 1586–1589, 2015.

APPENDIX

This appendix presents a list with all news sources that were used in the experiments to confirm the concept drift detection. These websites were last visited on March 19th, 2018.

dilmabr	
	http://varelanoticias.com.br/no-twitter-ex-presidente-dilma-rousseff-aponta-machismo-durante-sua-gestao https://www.metropoles.com/brasil/no-twitter-dilma-e-chamada-de-bandida-e-faz-piada-com-seguidor https://www.brasil247.com/pt/247/poder/325767/Dilma-posta-desabafo-nas-redes-contra-a-misoginia-dos-que-a-golpearam.htm http://diaonline.com.br/2017/11/05/dilma-e-chamada-de-bandida-e-faz-piada
LulapeloBrasil	
	http://noblat.oglobo.globo.com/artigos/noticia/2017/07/lula-e-o-chefe.html http://juntospelobrasil.com/joesley-relata-sobre-repasse-de-r-300-milhoes-a-lula https://www.oantagonista.com/brasil/imagine-se-lula-for-preso https://jornalggn.com.br/noticia/testemunhas-dizem-a-moro-que-lula-recusou-imovel-da-odebrecht
MichelTemer	
	http://g1.globo.com/economia/blog/thais-heredia/post/virada-liberal-um-ano-de-temer.html https://oglobo.globo.com/brasil/temer-recebeu-15-milhoes-para-aliados-roubou-1-milhao-para-ele-diz-diretor-da-jbs-21364963 http://www1.folha.uol.com.br/poder/2017/05/1888169-em-meio-a-crise-temer-troca-o-comando-do-ministerio-da-justica.shtml http://www.bbc.com/portuguese/brasil-40539366 https://exame.abril.com.br/brasil/as-duvidas-da-reforma-trabalhista https://g1.globo.com/politica/blog/andreia-sadi/post/2017/12/04/partidos-articulam-com-temer-pacto-para-barrar-em-2018-deputados-que-votarem-contra-a-reforma-da-previdencia.ghtml http://ultimosegundo.ig.com.br/politica/2017-12-24/michel-temer-pronunciamento-natal.html https://veja.abril.com.br/brasil/marun-admite-que-governo-negocia-credito-da-caixa-por-votos https://g1.globo.com/politica/noticia/votacao-denuncia-temer-camara.ghtml
ForaTemer	
	http://blogs.correiobraziliense.com.br/cbpoder/fora-temer-no-museu-nacional https://oglobo.globo.com/brasil/pintor-escreve-fora-temer-em-monumento-de-brasilia-21167559 http://www.diariodocentrodomundo.com.br/michel-teria-muito-a-aprender-com-o-heroi-anonimo-que-pintou-fora-temer-no-museu-nacional-de-brasilia-por-kiko-nogueira https://www.brasil247.com/pt/247/brasilia247/288919/Funcion\%C3\%A1rio-pinta-\%E2\%80\%9CFora-Temer\%E2\%80\%9D-no-Museu-Nacional.htm https://www.revistaforum.com.br/2017/04/05/funcionario-de-empresa-de-revitalizacao-pinta-fora-temer-no-museu-nacional http://politica.estadao.com.br/blogs/coluna-do-estadao/entrevista-fora-temer-so-beneficia-o-pt-diz-ministro-aloysio-nunes https://www.brasil247.com/pt/247/mundo/305103/Temer-\%C3\%A9-recebido-com-Fora-Temer-em-Hamburgo.htm http://economia.ig.com.br/2017-07-07/promocao-fora-temer-manifestacao.html http://www.diariodocentrodomundo.com.br/essencial/faixa-fora-temer-e-estendida-em-edificio-de-hamburgo-sede-do-g-20 http://politica.estadao.com.br/noticias/geral,oposicao-entra-em-plenario-com-cartaz-fora-temer,70001920779 https://noticias.uol.com.br/ultimas-noticias/agencia-estado/2017/08/02/partidos-da-oposicao-entram-no-plenario-da-camara-com-cartazes-de-fora-temer.htm https://www.brasil247.com/pt/247/brasil/309560/Marina-descobriu-o-Fora-Temer-diz-Fernando-Brito.htm http://www1.folha.uol.com.br/ilustrada/2017/09/1917344-gil-faz-show-com-clima-de-festa-em-familia-e-diz-que-temer-ja-esta-fora.shtml

petrobras	
	http://politica.estadao.com.br/blogs/fausto-macedo/pf-cumpre-10-mandados-em-nova-operacao https://oglobo.globo.com/brasil/lava-jato-pf-prende-ex-gerente-da-petrobras-por-destruicao-de-provas-21970481 https://www.istoedinheiro.com.br/petrobras-acoes-detidas-na-braskem-e-na-deten-estao-livres-para-venda http://www.fundacaoastrojildo.com.br/2015/2017/10/20/miriam-leitao-o-sinal-da-petrobras http://www1.folha.uol.com.br/poder/2017/10/1928660-pf-cumpre-mandados-em-operacao-sobre-desvios-na-petrobras.shtml http://www.valor.com.br/politica/5162928/pf-apura-pagamentos-de-vantagens-indevidas-petrobras http://agenciabrasil.ebc.com.br/geral/noticia/2017-10/pf-investiga-contratos-suspeitos-envolvendo-odebrecht-e-petrobras http://agenciabrasil.ebc.com.br/politica/noticia/2017-10/lava-jato-investiga-uso-de-subsidiarias-da-petrobras-para-favorecer https://www.em.com.br/app/noticia/politica/2017/10/20/interna_politica,910202/procuradoria-aponta-o-custo-da-corrupcao-na-petrobras.shtml
SenadoFederal	
	https://www12.senado.leg.br/noticias/materias/2017/07/04/senado-vota-reforma-trabalhista-incentivos-e-precatorios-antes-do-recesso http://www.jb.com.br/pais/noticias/2017/07/05/aecio-deve-desculpas-ao-senado-federal-e-ao-pais-diz-lindbergh-farias https://www.cartacapital.com.br/politica/plenario-aprova-regime-de-urgencia-para-a-reforma-trabalhista http://alec.org.br/novo/noticias/reforma-trabalhista-e-aprovada-na-ccj-do-senado-federal-e-segue-para-votacao-no-plenario-da-casa https://paraibaonline.com.br/2017/12/caciques-do-senado-com-reeleicao-complicada https://www.correiodoestado.com.br/politica/caciques-do-senado-na-mira-da-lava-jato-terao-reeleicao-dificil/318187 https://www12.senado.leg.br/noticias/audios/2017/12/senadores-comentam-liminar-de-lewandowski-que-garante-aumento-dos-servidores http://www.folhapolitica.org/2017/12/peticao-pelo-impeachment-de-gilmar_23.html https://www12.senado.leg.br/noticias/materias/2017/12/26/pec-exige-maioria-do-stf-para-afastamento-de-chefes-do-executivo-e-legislativo https://www.ignews.com.br/2017/12/26/pedido-de-impeachment-de-gilmar-mendes-acaba-de-ser-protocolado-no-senado-federal http://www.noticiatododia.com.br/2017/12/26/pedido-de-impeachment-de-gilmar-mendes-acaba-de-ser-protocolado-no-senado-federal
CamaraDeputados	
	http://www.gazetadopovo.com.br/politica/republica/entenda-como-sera-a-votacao-na-camara-da-denuncia-que-tira-o-sono-de-temer-4991vjie36cp61zz42kiqx8tw http://politica.estadao.com.br/noticias/geral,temer-recompoe-apoio-para-barrar-denuncia-na-camara,70001913431 http://agenciabrasil.ebc.com.br/politica/noticia/2017-07/camara-reinicia-sessoes-na-3afeira-e-pode-votar-reformas-da-previdencia-e https://g1.globo.com/politica/noticia/orientacao-e-marcar-presenca-em-plenario-diz-vice-lider-do-governo-sobre-votacao-de-denuncia.ghtml http://radios.ebc.com.br/reporter-nacional/2017/11/edson-fachin-desmembra-processo-que-investiga-organizacao-criminosa-no http://www2.camara.leg.br/camaranoticias/noticias/EDUCACAO-E-CULTURA/551156-CAMARA-APROVA-QUATRO-PROJETOS-QUE-CRIAM-UNIVERSIDADES-FEDERAIS.html http://www2.camara.leg.br/camaranoticias/noticias/POLITICA/551177-APROVADA-URGENCIA-PARA-PROJETO-QUE-REGULAMENTA-ATIVIDADE-DE-LOBBY.html https://noticias.uol.com.br/politica/ultimas-noticias/2017/08/25/com-impasse-sobre-fundo-e-distritao-camara-pode-aprovar-so-fim-de-coligacoes.htm http://www2.camara.leg.br/camaranoticias/noticias/DIREITOS-HUMANOS/541803-EVENTOS-NA-CAMARA-DOS-DEPUTADOS-LEMBRAM-OS-38-ANOS-DA-LEI-DE-ANISTIA.html

STF_oficial	
	https://veja.abril.com.br/brasil/abaixo-assinado-que-impeachment-de-gilmar-tem-16-mi-assinaturas https://www.ocafezinho.com/2017/12/22/procurador-da-lava-jato-ataca-presidente-da-republica-no-facebook http://epoca.globo.com/politica/expresso/noticia/2017/12/procurador-que-atuou-no-impeachment-de-dilma-critica-alguns-ministros-do-stf.html http://jovempan.uol.com.br/programas/jornal-da-manha/questionado-sobre-decisao-de-lewandowski-barroso-diz-que-stf-nao-deve-interferir-na-economia.html http://www.eb.mil.br/web/resenha/display/-/asset_publisher/9B8IpAnDp1we/content/pedido-de-maluf-e-negado-no-stf
politica	
	http://www.correiobraziliense.com.br/app/noticia/economia/2017/06/04/internas_economia,600019/crise-politica-poe-em-risco-inicio-da-retomada.shtml https://oglobo.globo.com/economia/crise-politica-gera-incerteza-para-aereas-no-brasil-diz-executivo-da-iata-21435583 https://brasil.elpais.com/brasil/2017/06/02/politica/1496433392_289468.html http://politica.estadao.com.br/noticias/geral,a-dilmizacao-de-temer,70001825165 http://g1.globo.com/jornal-nacional/noticia/2017/08/comissao-da-reforma-politica-tem-alternativa-para-fundo-de-campanha.html http://www1.folha.uol.com.br/poder/2017/08/1909299-reforma-politica-ou-conluio-contra-o-cidadao-brasileiro.shtml http://www1.folha.uol.com.br/poder/2017/08/1909503-dilma-critica-reforma-politica-em-velorio-do-ex-marido.shtml http://agenciabrasil.ebc.com.br/politica/noticia/2017-08/entenda-o-que-e-o-modelo-distritao http://www.correiobraziliense.com.br/app/noticia/politica/2017/08/23/internas_polbraeco,620182/manifestantes-invadem-comissao-da-reforma-politica-e-suspendem-sessao.shtml http://ultimosegundo.ig.com.br/colunas/antonio-riccitelli/2017-08-23/reforma-politica-brasil.html https://noticias.r7.com/brasil/camara-decide-fatiar-votacao-da-reforma-politica-23082017 http://politica.estadao.com.br/noticias/geral,para-metade-dos-brasileiros-saida-de-temer-nao-resolvera-crise-politica,70002007422 http://diariodonordeste.verdesmares.com.br/cadernos/nacional/online/brasil-precisa-desesperadamente-da-reforma-politica-diz-barroso-1.1822724 https://exame.abril.com.br/brasil/a-opiniao-do-brasileiro-sobre-lava-jato-crise-e-reforma-politica http://www.correiobraziliense.com.br/app/noticia/politica/2017/12/13/internas_polbraeco,647797/temer-passa-por-cirurgia-em-sao-paulo.shtml http://politica.estadao.com.br/noticias/geral,filha-de-roberto-jefferson-e-a-nova-ministra-do-trabalho,70002137395
dolar	
	http://www.correiobraziliense.com.br/app/noticia/economia/2017/06/19/internas_economia,603289/taxas-futuras-de-juros-sobem-com-dolar-de-olho-na-crise-politica.shtml https://g1.globo.com/economia/mercados/noticia/cotacao-do-dolar-190617.ghtml https://www.noticiasagricolas.com.br/noticias/politica-economia/193735-dolar-sobe-e-vai-a-r-331-com-apreensao-diante-da-cena-politica.html#.WmccTxYfMOI https://www.terra.com.br/economia/dolar-passa-por-correcao-e-cai-ante-real-monitorando-negociacoes-para-andamento-das-reformas,6bac2b09ae4fe815fb576196236d6e1e8koa3iow.html http://www.valor.com.br/financas/5086198/dolar-cai-em-semana-mais-volatil-desde-agravamento-da-crise-politica http://www1.folha.uol.com.br/mercado/2017/08/1911069-petrobras-puxa-alta-de-1-da-bolsa-dolar-corrige-e-cai-para-r-314.shtml https://economia.uol.com.br/cotacoes/noticias/redacao/2017/10/16/dolar.htm http://www.valor.com.br/financas/5156932/dolar-e-juros-futuros-sobem-de-olho-em-fed-e-cautela-na-politica https://g1.globo.com/economia/mercados/noticia/cotacao-do-dolar-161017.ghtml

dengue	
	https://noticias.r7.com/saude/veja-as-diferencas-entre-os-sintomas-da-chikungunya-e-o-da-dengue-29052017 http://blogs.ne10.uol.com.br/casasaudavel/2017/05/28/responsavel-por-epidemias-de-dengue-hemorragica-sorotipo-2-da-doenca-preocupa-em-pe https://www.reuters.com/article/us-sri-lanka-disaster-diseases/sri-lankas-flood-survivors-t\%hreatened-by-dengue-disease-aid-workers-idUSKBN18P1Q0 http://www.fiepi.com.br/sesi/noticias/geral/campanha-de-combate-ao-mosquito-da-dengue\%-prossegue-nas-ruas-e-instituicoes-publicas-4821.html https://veja.abril.com.br/saude/pernil\%ongo-pode-transmitir-o-virus-zika-revela-novo-estudo https://ricmais.com.br/sc/programas/balanco-geral-chapeco/atenc\%-chapeco-tem\%-439-focos-do-mosquito-da-dengue-em-2017 https://rededengue.fiocruz.br/numeros/626-monitoramento-dos-casos-de-dengue-c\%hikungunya-e-zika-ate-a-semana-epidemiologica-35-2017 http://hojeemdia.com.br/horizontes/capital-registra-802-casos-de-dengue-70-d\%e-febre-chik\%ungunya-e-17-de-zika-v\%C3\%ADrus-1.563047 https://globoplay.globo.com/v/6300560 https://www.revide.com.br/noticias/saude/casos-de-dengue-tem-alta-de-408-em-ribeirao-preto10 https://ricmais.com.br/sc/programas/balanco-geral-joinville/joinville-registra-243-focos-na-cidade-em-2017 http://portalitajuba.com/noticia/1464/prefeitura-realizou-dia-d-de-combate-a-dengue-na-praca-theodomiro-santiago http://www.ceunes.ufes.br/conteudo/dia-nacional-de-combate-dengue http://www.radiorural.com.br/noticias/27142-concordia-iguala-a-2016-e-ja-tem-17-focos-do-mosquito-da-dengue
zika	
	http://www.business-standard.com/article/current-affairs/spanish-scientists-discover-molecule-for-potential-drug-to-cure-zika-virus-117072300025_1.html http://www.nbc-2.com/story/36609915/why-you-should-still-be-mindful-of-the-zika-virus https://www.opovo.com.br/jornal/cienciaesaude/2017/09/virus-zika-pode-vir-a-ser-usado-contra-tumor-cerebral.html https://g1.globo.com/sc/santa-catarina/noticia/em-sc-numero-de-focos-do-aedes-aegypti-cresce-43-em-2017.ghtml http://jconline.ne10.uol.com.br/canal/mundo/brasil/noticia/2017/10/10/casos-de-malaria-voltam-a-crescer-no-brasil-mostra-ministerio-da-saude-310956.php https://www.neurologyad\%visor.com/general-neurol\%ogy/first-test-to-detect-zika-in-blood-donations-approved/article/698819 http://hojeemdia.com.br/horizontes/casos-de-den\%gue-em-bh-ultrapassam-os-800-conforme-secretaria-de-sa\%C3\%BAde-1.568192 https://g1.globo.com/am/amazonas/noticia/fmt-s\e\%leciona-pacientes-do-am-para-estudo-sobre-virus-da-zika.ghtml
saude	
	https://www.revide.com.br/noticias/saude/sus-recebe-medicamento-para-tratamento-de-microcefalia-09 https://www.oncologiabrasil.com.br/ministerio-da-saude-incorpora-pertuzumabe-ao-sus-para-o-tratamento-do-cancer-de-mama https://www.correiodoestado.com.br/ciencia-e-saude/proposta-autoriza-medicamentos-nao-incluidos-na-lista-do-sus/316942 http://www.douradosagora.com.br/noticias/ciencia-saude/operadoras-de-planos-de-saude-repassaram-r-459-milhoes-ao-sus-ate-outubro http://diariodonordeste.verdesmares.com.br/cadernos/negocios/online/ans-repassa-valor-recorde-ao-sus-em-2017-1.1869069 http://www.sindessmat.com.br/portal/index.php/2014-03-08-20-06-44/item/5828-ans-repassa-valor-recorde-ao-sus-em-2017