



Enhancing decision-making support by mining social media data with social network analysis

Manuela Freire^{1,2} · Francisco Antunes^{3,4} · João Paulo Costa¹

Received: 2 February 2023 / Revised: 26 April 2023 / Accepted: 27 April 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

Abstract

This paper explores the use of social network analysis (SNA) on airlines' online social networks (OSNs) to extract valuable information for decision support, by analyzing interactions and discursive exchanges between users. The research is focused on fostering customer service of an airline company during a strike period, namely by detecting influential customers (whether satisfied or dissatisfied), address pending requests, and enhancing customer satisfaction, thus promoting issue-solving, and increasing responsiveness. The methodology involves analyzing data from the Facebook account of an airline company, using SNA to structure the data, and calculating metrics to detect possible situations to be addressed by customer service. The research concludes that it is possible to extract valuable information for decision support by analyzing the metrics that were built over the interactions and discursive exchanges between OSN users. SNA metrics enable to measure airline's call-center performance in terms of speed of answer and customer satisfaction, to identify active users requiring additional support, as well as highly influential customers who may impact on the overall customer satisfaction, thus helping to resolve issues more efficiently. This study provides both theoretical and practical implications: it contributes to the existing literature by integrating social interaction and SNA for decision support in airline's service context; and it provides practical insights into how companies can use SNA metrics to improve customer service. The research also highlights and corroborates the importance of monitoring social media interactions for decision-making and improving customer service.

Keywords Social network analysis · SNA · Data mining · Decision · Discourse · Network analysis

1 Introduction

The evolution of online social networks (OSN) has a crucial role for the rapid dissemination of information. Information and communication technologies have increased, almost

unlimitedly, the capacity to communicate, to inform and to be informed, as well as to acquire knowledge (Fu et al. 2017; Savic et al. 2019) from interpersonal networks that are always producing information (Piedrahita et al. 2017). It is recognized that digital is disrupting business models, while creating new opportunities for companies to engage, not only with their customers, but with all stakeholders.

This paper discusses how social network analysis (SNA) can help to get useful information from network structures and relationships to support decisions. To develop decision-support from social interactions, we analyzed data from a Facebook fan page of a well-known Portuguese airline (TAP) account, during a ten-day period strike and during the following two days. The main goal was to assess the airline's service level and responsiveness, as well as the development of business aids to evaluate and/or develop strategies to improve customer service, during a strike, based on the information retrieved from social media. The case study demonstrated that SNA metrics such as degree, closeness, and betweenness centrality can be used to identify influential

✉ Manuela Freire
maria.s.freire@fe.uc.pt

Francisco Antunes
francisco.antunes@ubi.pt

João Paulo Costa
jpaulo@fe.uc.pt

¹ Faculty of Economics, CeBER, Univ Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

² INESCC - Computer and Systems Engineering Institute of Coimbra, Coimbra, Portugal

³ Department of Management and Economics, University of Beira Interior, Estrada do Sineiro S/N, 6200-209 Covilhã, Portugal

⁴ INESCC - Computer and Systems Engineering Institute of Coimbra, Coimbra, Portugal

customers, improve response times, identify active users in need of additional support, and resolve issues more effectively. Additionally, SNA metrics such as in-degree, out-degree, PageRank, and eigenvector centrality can be used to measure the airline's performance in terms of speed of answer and customer satisfaction, helping call-center managers to identify the most active users in the network, the need for additional support, and highly influential customers who may impact overall customer satisfaction.

OSNs can be monitored to filter social conversations (Freire et al. 2022), providing businesses tools to better understand their customers and to build customer profiles. Moreover, by combining different techniques it is possible to interpret social networks' activity. Computerized systems can capture these social conversations, based on customer needs and their interactions. Together, with SNA, data mining (DM) and natural language processing (NLP) algorithms for OSN, content analysis can be used to interpret web discourse. In fact, in business contexts, it is important to understand customers (Isson 2018; Liu et al. 2020; Wang et al. 2020), as quickly as possible, to improve decision making.

SNA incorporates several centrality metrics that aid in studying social networks (further details will be provided along Sect. 2). NLP is the computer science field in which machines extract knowledge from natural language texts (Antunes et al. 2016; Isson 2018) and DM regards the automated retrieval of noteworthy data structures, integrating data analysis to infer models to meet specific objectives (Fu et al. 2017; Yang et al. 2022).

The extraction of predictive information, using the interactions between OSN's users, is something to explore. Understanding relationships between messages and interpreting the meaning of what is being exchanged usually relates to speech analysis, a multidisciplinary area itself. Furthermore, web discourse is linked to technologies in which communication is characterized by a fusion between written language and spoken discourse.

As web discourse often encompasses knowledge construction, the interaction, and the consequent development of OSN's, for capturing and understanding all interactions, can be extremely valuable in developing decision support. Our work determines that building upon social interactions it is possible to develop such support, as the acquisition, analysis, treatment, and use of social information do generate insights for decision-making.

After COVID-19 pandemic, the world has been watching a travelling restart, but not without many lockdowns and/or strikes, especially in civil aviation. Sources as Bartos and Badanik (2019) and SimpliFlying (2019), establish that the attention to social media is now part of airlines' daily operations, namely to bond with passengers. Despite this situation, social network analysis is not a mainstream to support decision-making in crisis management, though SNA can

provide a deep understanding of online social interactions (Freire et al. 2022). Moreover, we stand that such information is not only easy to obtain, but that it is a valuable cost-efficient resource because there are many free-available tools that can produce the information, as presented in this paper.

In Sect. 2 we introduce concepts on decision-making using SNA and we present an overview of studies regarding OSN's and airline services. Section 3 covers web discourse in OSN's context. In Sect. 4 we present a case study, and the obtained results are addressed in the following section. Final remarks are presented in the last section.

2 Social network analysis and decision-making support for airline services

Making decisions often involves uncertainty. This can be due to the fact that decision-makers may plan actions that will take place in the future or because their decisions are not always based on actual knowledge or absolute principles, but on assumptions, perceptions, opinions, and personal values or beliefs. Fedushko et al. (2023) suggest that strategic decision-making, based on OSN, involves understanding the interconnected decisions and their impact on network users to gain an advantage. On the other hand, Antunes and Costa (2011) argue that decision-making support processes should incorporate rational aids instead of mere "gut feelings", thus enhancing the possibility of making better decisions. Freire et al. (2022) build upon these ideas to establish that decision-making processes dependent on OSN could benefit from the use of SNA metrics. These metrics can help to understand the interaction of users, involved posts, and encompassed concepts, providing a deeper understanding of the discussion and its logical sequence.

Social network data and information are usually associated with computer systems. However, OSNs are considered an essential source of real-time information for decision-making (Arafah et al. 2021) and it might be a manager's greatest technological innovation for decision support. As OSNs involve virtual collaboration, including the execution of virtual organizational tasks, understanding the effects of online collaboration and information sharing is crucial for supporting decision-making in the current context, when compared to face-to-face collaboration. In contemporary organizations, decision-makers need to gather relevant information to increase their knowledge of available alternatives for a particular choice, by means of analysis tools, in order to capture and analyze everything a social user publishes (texts, images, etc.), to make faster judgments on the obtained results.

A social network facilitates the exchange of information and knowledge (Wu et al. 2017; Camacho et al. 2020;

Aggrawal and Anand (2022) by connecting various entities through social relationships (Wasserman and Faust 1994; St-Onge et al. 2022). To understand the impact of this kind of network on group behavior and individual influence, SNA can be used to describe its structure (Wasserman and Faust 1994). Studies on web discourse primarily focus on computational aspects (Herring 2013), such as NLP (Yuksel and Tan 2018; Giarelis et al. 2021), online content analysis, and speech analysis (Moser et al. 2013; Kok and Rogers 2017; Pennington 2017; Goritz et al. 2019), with a significant emphasis on text analysis and utilizing other knowledge areas. However, SNA and semantic analysis of posts have not received much attention and remain largely theoretical (Zhao et al. 2021). Freire et al. (2022) provide an overview of such studies and conclude that they mostly focus on rhetoric, argumentation, semantics, and pragmatic issues rather than discursive aspects for text analysis on the web. Text classification algorithms, such as Word2Vec proposed by Mikolov et al. (2013) or GloVe (Pennington et al. 2014), have been developed to represent semantic information from a given content. Alam et al. (2020) explore the use of NLP and computerized visualization techniques to leverage information created during disaster events and suggest combining different computational techniques in such contexts. Fani et al. (2020) use temporal content similarity to learn neural embeddings for users and neural graph embedding techniques to use social network structures to represent information. Wankhade et al. (2022) discuss various methods to obtain and analyze people's opinions on different topics, as well as the application of sentiment analysis.

Despite the existence of some theoretical frameworks, most previous studies on airlines' OSNs have been exploratory and mainly focused on machine learning or lexicon-based methods to categorize airline companies' OSN content. Many of these studies only regarded the application of sentiment analysis to textual content. For example, Dutta Das et al. (2017) analyzed Twitter data from Emirates and Jet Airways using the Naïve Bayes algorithm in R and Rapid Miner tools to map tweets into three categories (positive, neutral, and negative). Rane and Kumar (2018) used pre-processing techniques to clean tweets from six US airline companies and used Doc2vec, a concept of deep learning, to perform phrase-level analysis that takes word order into consideration, leading to an aggregation of sentiments in three categories using different classification strategies. Tiwari et al. (2019) presented positive and negative sentiments and their correlation regarding customer tweets, by means of BIRCH (balanced iterative reducing and clustering using hierarchies), a hierarchical clustering approach, combined with rule mining techniques. Bartos and Badanik (2019) emphasized the importance of OSNs for airlines and presented examples of how airlines can deal with communication crises, contributing to airlines' revenue growth. Naseem

et al. (2019) presented a deep intelligent contextual embedding to solve language complexities and embed word sentiments. The authors' model addresses the complexities of other languages with a character embedding. Al-deen et al. (2021) proposed a combination of deep neural networks with a multi-head attention mechanism. They also conducted sentiment analysis and text classification tasks using a large Twitter dataset on the major US airline problems to experiment the performance of their model. Kumari and Prajna (2021) used a Twitter dataset of passengers' feedback of a US airline to propose a combination of classification methods and bagging of machine learning-based text classification techniques for the development of a voting classifier.

As none of the previous studies have considered the integration of all network elements, web discourse entities, process steps, or SNA techniques, in this study we present such integration by analyzing an airline's OSN, using SNA metrics, that act as indicators for possible customer service actions. To make this article more self-contained we present, in Table 1, a brief explanation of the used measures to determine the centrality, influence, and similarity of posts (from which we determined the conclusions that will be presented in Sect. 5).

3 Web discourse in OSN's context

OSN's comprise interacting processes in which emerging networks are built through an established dialogue (discourse). The complexity associated with the analysis of social relations, social interaction data analysis requires the use of methods and techniques, studied and validated in different areas of knowledge, to explain the discursive phenomena of communication. The language used in OSN's tends to be casual and informal, mixing linguistic and non-linguistic resources (Freire et al. 2015), especially to communicate quicker and express negative emotions, and, in the airline context, it also bares a multi-lingual nature (Antunes et al. 2014). This usually produces tangled, disordered, incomplete, and sometimes garbled data.

Briefly, the discourse produced on Facebook has a minimum of 3 entities (user, post, and concept) that are represented as square matrices, while adjacency and affiliation matrices are used to depict web discourse. The adjacency matrix identifies and formally represents the links between a single entity and the affiliation matrices characterize the links between each of them. This type of affiliation matrices (Wasserman and Faust 1994; Borgatti 2009) can be used to analyze interactions among entities at three different levels, namely the interactions of: users, users and posts, as well as posts and concepts. Using information from our case study, Fig. 1 exemplifies the matrices of the three web discourse entities.

In order to provide a formalization to analyze an airline's Facebook content and to determine its service level and overall responsiveness, two stages are needed: firstly, the analysis of the semantic contents embedded in

social web discourse and, secondly, the social interactions between users, e.g., the links embedded within the social network structure. We distinguish the entities and

Table 1 SNA metrics

SNA metrics	Description
Degree centrality	Represents the number of links that a node has within an undirected network. In other words, nodes that have many neighbors are considered central. In directed networks, the analysis of degree centrality is broken down into two categories: in-degree centrality and out-degree centrality. In-degree centrality is used to represent nodes that receive a higher number of interactions within the network, while out-degree centrality encompasses nodes that have a higher number of interactions throughout the network (Burguillo 2018; Savic et al. 2019)
Closeness centrality	Indicates the distance of a node to all others, and it can show participation differences, as it evidences node independence, revealing which users communicates to other with lesser intermediaries (Golbeck, 2015; Qasem et al., 2017; Burguillo 2018; Koutra and Faloutsos, 2018; Ishfaq et al., 2021; Aggrawal and Anand 2022)
Betweenness centrality	Indicates whether a node acts as a bridge among other nodes, meaning that high betweenness values reveal nodes that are central in a network, contributing for information flow and network cohesiveness (Wasserman and Faust 1994; Chaabani and Akaichi, 2022; Freire et al. 2022)
Eigenvector	Measures the influence of nodes, assigning scores to all network nodes, knowing that connections to high-scoring nodes have a larger contribution to a node's score, than the ones provided by low-scoring nodes (Zheng and Skillicorn, 2017; Hansen et al. 2020; Aggrawal and Anand 2022)
PageRank	Is a metric that assigns a score to network vertices based on the quantity and quality of links that point to them, as a way to determine their significance and importance in the network (Page et al., 1999; Fu et al. 2017; Aggrawal and Anand 2022; Bautista and Latapy, 2022). PageRank assigns a score to network vertices, to indicate their significance (Aggrawal and Anand 2022). In this metric, the importance of a node is determined by the quantity/quality of links that point to it, measuring the importance of a node (Fu et al. 2017)

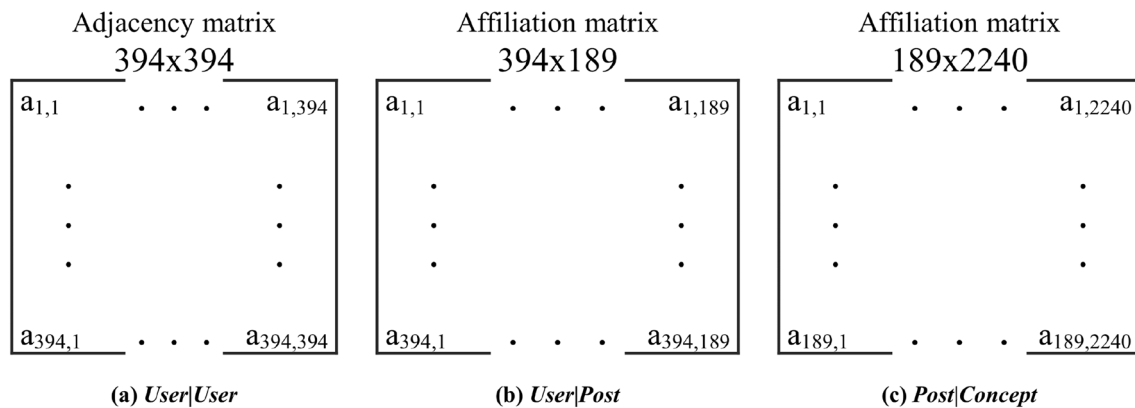


Fig. 1 Adjacency and affiliation matrices of the web discourse

Table 2 Type of entities and connections

Type	Description
User	Represents all users (customers or not) that engage to the airline's OSN. They are defined by their reciprocal relationships on the network and the reactions they provide to more than one post
Post	Represents the content that is shared by users in the airline's OSN. A post could be any searchable entity
Concept	Represents the textual data embedded in one, or more, post or comment reply
Interaction	Represents sets of interconnected users that can be linked using different criteria, such as friendship, behavior, or just belonging to the airline's OSN
Reaction	Represents user feedback

connections described in Table 2, as main components of the information provided by OSN, as depicted in Fig. 2.

According to graph theory (Wasserman and Faust 1994; Borgatti 2009; Burguillo 2018; Bechberger and Perryman 2020; Galety et al. 2022), instead of using the three matrices (adjacency and affiliation matrices) it is possible to represent a network composed of several entities as a two-mode network. Therefore, in order to create the semantic network embedded within the web discourse of the airline's OSN, we defined, in Eq. 1, the web discourse graph (WDG) as a tuple:

$$\text{WDG} = (U, P, C, L_1, L_2, L_3, L_4) \quad (1)$$

where $U = (u_1, \dots, u_k)$ is the set of users, $P = (p_1, \dots, p_w)$ was the collection of content produced by U , and $C = (c_1, \dots, c_m)$ encompassed all the concepts within a post or a comment. In Eq. 2, we represented the concept $c \in C$ of the produced content post $p \in P$ of each user $u \in U$ as:

$$X_{\text{upc}} = [x_{\text{upc } 1} \cdot x_{\text{upc } n}] \quad (2)$$

which we referred to as the user's produced content, where $x_{\text{upc } n} \in \mathbb{R}[0,1]$ shows the concept c of a post p produced by a user u . The equation represents the vector of user u 's produced content for a particular concept c in the WDG. Each element of the vector, $X_{\text{upc } n}$, represents the degree of association or relevance that the content created by user u has with concept c , where $X_{\text{upc } n}$ is a real number between 0 and 1. The vector X_{upc} is used to represent the user's contribution to the concept c in the OSN and to calculate various network measures, such as centrality, influence, or similarity, that help to analyze and understand the structure and dynamics of the WDG. L_1 , formalized in Eq. 3:

$$L_1 = (u_i, u_k) \in U \times U \quad (3)$$

refers to a set of links, that reflect each user u_i interaction with user u_k . The notation $U \times U$ indicates that the links are established between two elements of the set U , which represents all the users in the OSN. Specifically, the links in L_1 represent the interactions between users (such as direct messages, comments, or reactions to posts), which help to capture their social ties. The links in this equation are bidirectional or undirected, as directionality is not specified. L_2 , presented in Eq. 4 as:

$$L_2 = (u_k, p_w) \in U \times P \quad (4)$$

represents a set of links reflecting each comment of user u_k regarding a post p_w . The notation $U \times P$ indicates that the links are formed between an element of the set U , which represents all the users in the OSN, and an element of the set P , which encompasses all the content produced by users. Specifically, the links in L_2 represent comments posted by a user in response to a particular post, which helps to capture the interaction between users and the content in the OSN. The equation of L_3 , described in Eq. 5 as:

$$L_3 = (p_1, p_w) \in P \times P \quad (5)$$

integrates a set of links within the WDG, reflecting each comment p_1 posted in a post p_w (typically a comment reply). The notation $P \times P$ indicates that the links are generated between two elements of P , which encompasses all the content produced by users. Specifically, the links in L_3 represent comments posted in response to a particular post, forming a structure that is similar to a thread or discussion. Equation 6, which describes L_4 :

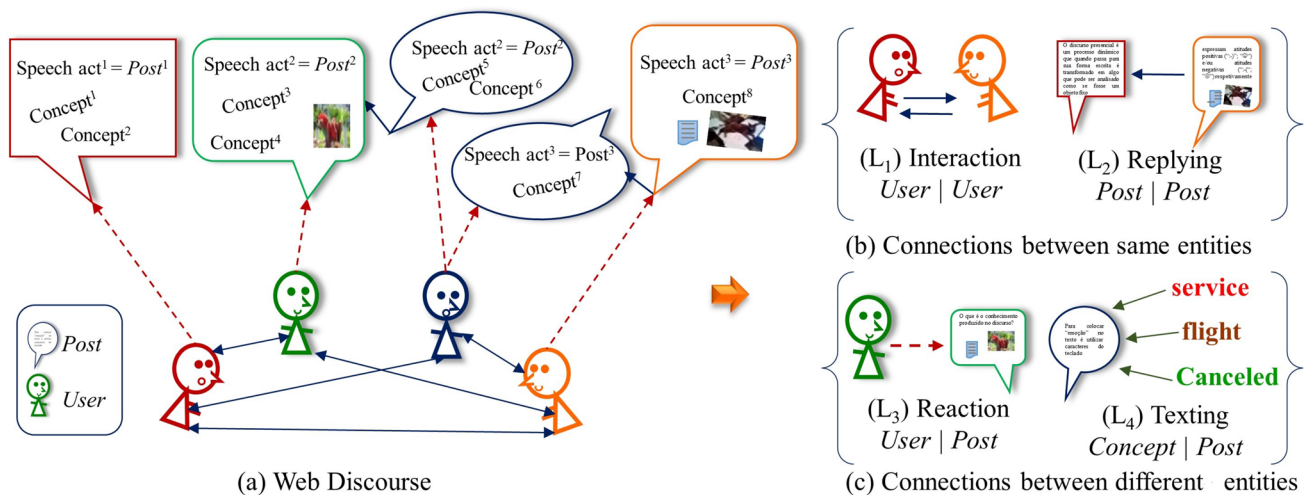


Fig. 2 Web discourse abstraction

$$L_4 = (u_k, p_w, c) U \times P \times C \quad (6)$$

is constituted by a set of links, that reflects each user u_k that reacted to post p_w , by using textual content. The notation $U \times P \times C$ indicates that the links are established between an element of the set U , which represents all the users in the OSN, an element of the set P , which encompasses all the content produced by users, and an element of the set C , which represents the concepts within the content. Explicitly, the links in L_4 represent a user's reaction to a post, using textual content that relates to a specific concept, allowing to capture how users engage with and interpret concepts within the content of the OSN. The global network of all users, posts, and concepts generates a bipartite graph of the WDG (described in Eq. 7):

$$X = [x_{upc} : u \in U, p \in P, c \in C] \quad (7)$$

where each node represents either a user, a post, or a concept, and the links represent the associations between these entities. The graph X is constructed by taking the cartesian product of the sets U , P , and C , and then for each combination of u , p , and c , creating a node with the vector of produced content X_{upc} that represents the degree of association or relevance between user u , post p , and concept c . Thus, the bipartite graph X captures the relationships between users, posts, and concepts, and is used to analyze the structure and properties of the OSN, such as clustering, community detection, or network flow.

To create an interaction network, we defined, in Eq. 8, the customer user social network structure as a directed graph:

$$G_{in} = (U, L) \quad (8)$$

where U represented the set of users, and L the connections between the users. These connections were ordered pairs of users $(u_j, u_k) \in L$ representing a social tie from u_j to u_k (e.g., u_j interacting with u_k). These connections were also ordered pairs of users and posts elements such as $(u_k, p_w) \in L$, representing a social reaction tie from u_k to p_w .

To create the semantic network, we defined, in Eq. 9, such network as a directed graph:

$$G_{sn} = (P, C, L) \quad (9)$$

where P was the content corpus (e.g., posts and comments), C encompassed the textual data, concepts, within posts, and L included the connections from the concept to the post. These connections were ordered pairs of concepts and posts elements such as $(c_k, p_w) \in L$ representing a reaction tie from c_k to p_w . To build the semantic network, we identified the set of valid concepts Z that were extracted, and we linked each content to each concept $z \in Z$, i.e., X_{upz} . Later, to extract the available concepts from the users' content collection and to link them to the posts, the cleaning algorithm, as well as

the cleaning database developed by Freire et al. (2022), was applied. We aggregated all posts and comments into a single dataset that produced our semantic corpus. We identified the most relevant post in P , by calculating in-degree, PageRank and eigenvector metrics, and the key concept in Z , by calculating the out-degree metric, for every two-day datasets (all 6 of them), as further explained in Sect. 5.1.

4 The case study

Following the framework in Freire et al. (2017), users' activities and supplementary data were collected, followed by data processing and representation. Data collection was carried out between May 1st and 12th. When comparing the various snapshots, it was observed that the first snapshot, referring to May 1st and 2nd, recorded the highest number of users (394). It was also observed that the data set for May 5th and 6th contained the highest number of posts and concepts (201 posts and 10,730 concepts, respectively). In total, 1328 users were identified, and 877 posts containing 43,969 concepts were collected.

A graph database was used to store structured data, while unstructured data were cleaned for a later standardization process before graph analysis and data visualization and the production of the output files. NodeXL (Hansen et al. 2020) was used for a periodical data extraction of two days, over the period of twelve days (creating six datasets, in which the first five regarded strike days and the last one an afterwards period). When collecting the data, only a set of OSN snapshots was permitted when using raw data from the user/user network, which led to the creation of two additional networks, to identify the relevance of posts, and to determine which concepts were used the most.

Using data mining and relational database techniques, as well as a common spreadsheet, we were able to aggregate the matrices regarding the three entities that support web discourse and represent the main network. Processed data were treated in Gephi because it not only responds to the needs of graphic studies using graphical user interface, but also due to its open source nature and free availability (Bastian et al. 2009), that allows the visualization and manipulation of the networks, if intended, and a later application of SNA techniques for its analysis. By using social network metrics, the most relevant nodes (users, posts, or concepts) were identified and classified according to their importance. For ethical, legal and privacy concerns, all user data was made anonymous, abiding national and international laws.

Semantic processing was also an important step of the analysis process, in which the data obtained from the retrieved 769 posts, was transformed into semantic networks. Contrasting to formal publications, the contents generated on OSN are not bound by grammatical and spelling rules, as

abbreviations are prone to happen in web discourse texting, especially within posts. To solve this problem, the framework proposed by (Freire et al. 2017) adopted text mining to discover posts' embedded concepts and stemming to normalize synonyms within posts, using a synonym table (for instance, every form of the concept "cancel", such as "cancelling", "canceled" and "cancelled", were converted to a single identical stem concept "cancel", thus helping to reduce the sparsity of the output dataset).

The cleaning data process that was configured for each lexical language of each post involved: (1) the removal of symbols, punctuation or spaces between concepts; (2) the search for emoticons and their replacement by a standardization term; (3) the search for punctuation characters, and also their replacement by a standardization term; (4) the search for stop words and their removal; (5) the normalization of synonyms to substitute the identical, or almost identical, meaning of concepts with a single one. The output posts encompassed 20,438 concepts in 7 different languages (German, Spanish, French, Italian, Dutch, Portuguese, and English).

5 Results and data visualization

After using Gephi to create the networks for each dataset (snapshot), we calculated the centrality measures of each dataset, as shown in Table 3.

As expressed in Table 3a, the airline company and fan page owner (TAP) exhibited the highest values for in-degree, out-degree and betweenness centrality. Such results tell that the airline company was quite active (regardless of some TV interviews, conducted at the time, that expressed the opposite) and did have some degree of control over the information flow.

Regarding closeness centrality, and as indicated in Table 3, many users (other than TAP) got the highest possible value (92 users in the first dataset; 83 in the second; 102 in the third; 58 in the fourth; 68 in the fifth; and 28 users in the last one), meaning not only that they communicated

much more directly (or with lesser intermediate users) to other users than the airline company did, but also that they were able to disseminate information throughout the network rather quickly, as closeness centrality is related to the amount of time in which a certain information gets shared by all network nodes. It is important to notice that the fact that none of those users were TAP does not imply that there was no closeness between TAP and other users of the page. It only indicates that, in each time interval, there were several other users who had a shorter path to other nodes than TAP did. In a strike situation like this, determining the users with higher closeness centrality seems rather important to mitigate users' flaming, eventually by addressing their complaints more thoroughly, kinder, or quicker, as their interventions will propagate faster throughout the network, generating a "fireball" effect (by flaming others). This sort of user management was not addressed by the airline company during the ten days of strike, as only in the latest dataset, when the strike was over, we observed a reduction of users with high closeness centrality.

According to PageRank and eigenvector metrics, the airline company got the higher values most of the time, which corroborated a control in information flow, and a centrality position within the network, although two other users gained relevance during the first half of the strike. When those users got their situation solved, the airline company became central within the network once again.

To analyze the messages, in-degree, PageRank, and eigenvector, were the selected metrics (as depicted in Table 3b), because in-degree quantifies incoming connections from users to posts, while PageRank and eigenvector accounted for the amount and quality of a node's connections and the centrality of its connected nodes, respectively. In Table 4 we summarize the obtained results, for each metric, when analyzing the 6 snapshots of the airline company.

Snapshot 3 was the period in which the highest values were recorded for in-degree (for posts or comments published on the fan page), out-degree (for airline replies) and betweenness centrality. If we think of Facebook as a call-center service, it can be assumed that in-degree quantified

Table 3 Most influential users and most relevant posts, according to SNA metrics

#	Snapshot (days)	(a) Most influential user						(b) Most relevant post		
		In-Degree	Out-Degree	Closeness	Betweenness	PageRank	Eigenvector	In-Degree	PageRank	Eigenvector
1	1st and 2nd of May	TAP	TAP	*	TAP	TAP	User-94	Post-35	Post-35	Post-1
2	3rd and 4th of May	TAP	TAP	*	TAP	User-58	TAP	Post-82	Post-82	Post-82
3	5th and 6th of May	TAP	TAP	*	TAP	User-79	User-79	Post-178	Post-178	Post-4
4	7th and 8th of May	TAP	TAP	*	TAP	TAP	TAP	Post-74	Post-74	Post-74
5	9th and 10th of May	TAP	TAP	*	TAP	TAP	TAP	Post-50	Post-50	Post-21
6	11th and 12th of May	TAP	TAP	*	TAP	TAP	TAP	Post-69	Post-69	Post-69

(*) It means that several users, apart from TAP, have received the highest possible value of 1.000

Table 4 SNA metrics for the airline company

#	Snapshot (days)	In-Degree	Out-Degree	Closeness	Betweenness	PageRank	Eigenvector
1	1st and 2nd of May	92	277	0.777	0.008	0.008	0.262
2	3rd and 4th of May	77	231	0.824	0.006	0.008	1.000
3	5th and 6th of May	130	369	0.887	0.010	0.009	0.680
4	7th and 8th of May	85	248	0.966	0.008	0.011	1.000
5	9th and 10th of May	51	162	0.833	0.006	0.007	1.000
6	11th and 12th of May	49	115	0.850	0.003	0.008	1.000

the flow of incoming calls, and that out-degree quantified the effort in answering the calls. From this perspective, it is possible to prove the effort from the airline company in responding to user requests. Although it is not possible to assure from data, whether such effort turned out to be a mere apology or an actual resolution of a problem, it was, nonetheless, an effort in reaching users, and providing some sort of feedback, as expected from customer service (an apology will always be a far better response than mere silence). In-degree, out-degree, and betweenness centrality, therefore, can be used not only as thresholds for customer-level standards, but also as verification metrics of the defined thresholds.

Still in Table 4, the fourth snapshot evidenced the highest value in closeness centrality and PageRank (0.966 and 0.011 respectively), meaning that those were the days when the airline company directly communicated with more users and that such activity was felt of actual importance. At this point in time, 6 days of strike had already elapsed, and the airline seemed to have picked up the pace and deployed more resources into getting to know users' problems and to address them, while reducing the pressure for information/contact from the airline company. Due to this fact, we stand that it is important to get closeness centrality and PageRank as high as possible, as it translates into a higher customer service level.

5.1 Web discourse network

Improved visualization and clearer presentation could greatly enhance numerous discursive exchanges on OSNs. Distinguishing topics, disagreements, and points of view more effectively, as well as navigating through critical questions rather than following a chronological order, are crucial needs that must be addressed. To analyze the OSN data and to display the final information to decision makers, which is presented in tables of a spreadsheet, several intermediary networks are seamlessly built. If intended, each of those networks could also be depicted for analysis. For academic purposes we will present the most illustrative ones. In Fig. 3 we evidence the web discourse network with all entities, while in Figs. 4 and 5 we depict the sub-networks user|post and post|concept, respectively.

In Fig. 3, from (a) to (f), we represented the obtained results for all the 6 datasets. Identified clusters use modularity class, nodes' size is based on betweenness centrality and labels' size uses in-degree. Larger nodes indicate that they are more central to the network, playing a critical role in information flow and network cohesiveness. The size of the labels assigned to each node is proportional to the number of interactions received by the corresponding user or post, providing valuable information about the most important entities in the network. This information can be used to identify the most commented or viewed posts, as well as the most influential users within the network. Similar colors represent different subgroups, as well as denser interconnected nodes. Each depicted network conveys users, posts, and concepts, and each one represents a single dataset of a two-day period, as explained before. All representations in Fig. 3 were obtained by applying the Network Splitter 3D visualization algorithm (Barão, 2014) in Gephi, after the application of ForceAtlas3D Layout visualization algorithm.

After determining the above-mentioned centrality metrics, all six networks were disaggregated and only two of the entities of the web discourse were analyzed: users and posts. As seen in Fig. 4, identified clusters use modularity class, nodes' size is based on in-degree centrality and labels' size uses out-degree. Moreover, the size of each node is proportional to its importance within the network. Larger nodes indicate users that are more central and receive a higher number of interactions within the network. The label size, as in Fig. 3, reflects the number of user interactions throughout the network, highlighting the most active users.

From the posts it was possible to determine that many users complained that they got no answer from the airline's call-center, nor any response to their emails, while service offices were closed. Additionally, the airline's website did not provide any information about the strike. Despite contrasting with the earlier obtained perspective from Table 4 (in which was proven that there was an effort from the airline company in responding to user requests), this situation can mean that not every user was, in fact, assisted. This is corroborated by the isolated nodes between users (and their posts) and the airline company (TAP) that are possible to perceive in Fig. 4, meaning that the airline did not address the posts from those people. Nonetheless, not every isolated

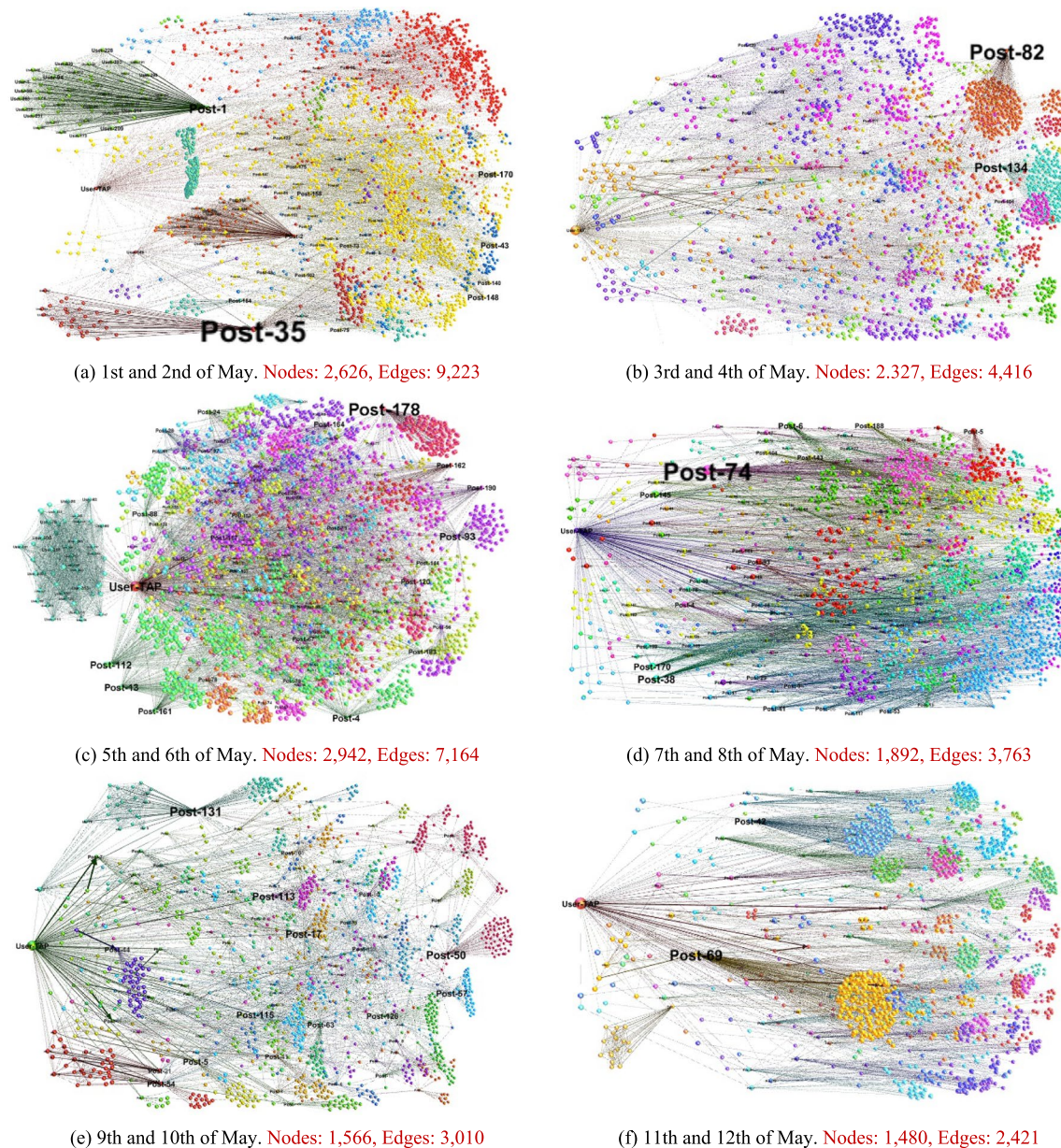


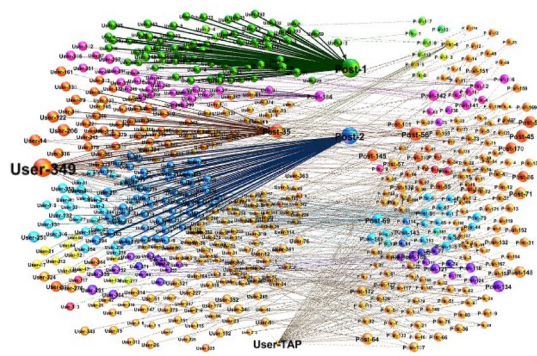
Fig. 3 Network: user|post|concept. **a**, **b**, **c**, **d** and **e** networks during the ten days of the strike. **f** network on the following two days. Each network consists of nodes representing users, posts, and concept entities, and edges representing connections between them. The nodes

are color-coded based on subgroups, and the size of each node represents its importance in the network (a more detailed visualization is available at https://drive.google.com/drive/folders/1cmrV2eksItWloBKVtB1hF2NXYAvStRw?usp=share_link)

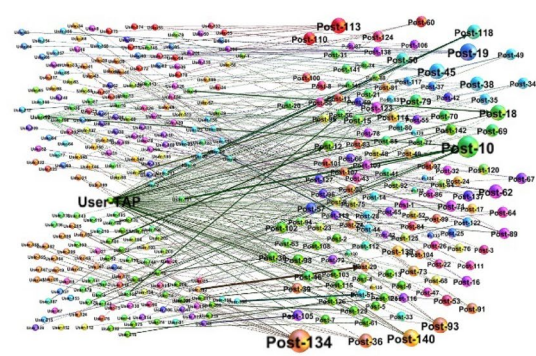
node meant an unaddressed request and taking a deeper look into Fig. 4a, we can observe, for instance, three large, isolated posts (a green, a blue, and an orange one) where no such thing happened. For instance, the green post (post-01) was a message posted by a user (User-94) informing that, due to the strike, he would not be able to perform that night in a well-known London electronic music club. This post got many other replies and comments (thus having the highest metrics, namely an in-degree of 66, a PageRank of 0.062 and an eigenvector of 1); the blue post (post-02) was a political

endorsement to the airline as a strategic asset for the country and the need to provide it with better resources to fully develop its potential, that also got many replies (achieving an in-degree of 56, a PageRank of 0.053 and an eigenvector of 0.848); and the orange post (post-35) was an open critic to pilots and their high wage level, that, again, got many replies (achieving an in-degree of 30, a PageRank of 0.026 and an eigenvector of 0.455).

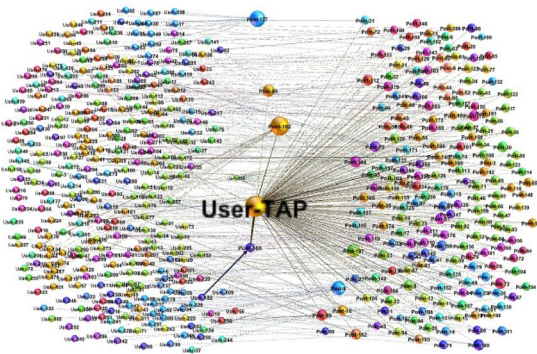
From an airline company point of view, identifying these posts (with high in-degree, PageRank, and eigenvector



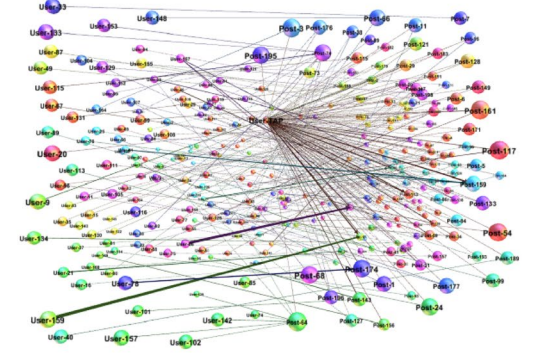
(a) 1st and 2nd of May. Nodes: 583, Edges: 614



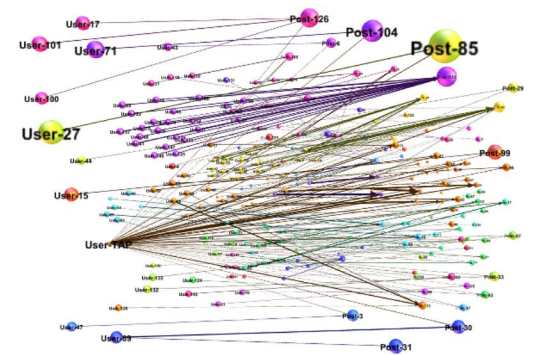
(b) 3rd and 4th of May. Nodes: 359, Edges: 360



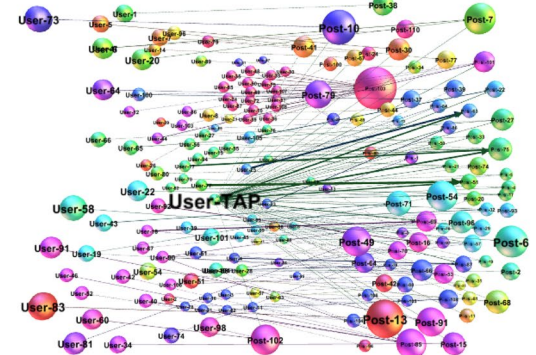
(c) 5th and 6th of May. Nodes: 517, Edges: 546



(d) 7th and 8th of May. Nodes: 312, Edges: 311



(e) 9th and 10th of May. Nodes: 251, Edges: 255



(f) 11th and 12th of May. Nodes: 184, Edges: 186

Fig. 4 Network: userpost. **a**, **b**, **c**, **d** and **e** networks during the ten days of the strike. **f** network on the following two days. Each network comprises nodes that represent users and posts, along with edges that depict connections between them. To distinguish different subgroups,

nodes are color-coded accordingly (a more detailed visualization is available at https://drive.google.com/drive/folders/1qIU0u3cXu-q0DfO-f1lrpMxGtezEN-E-F?usp=share_link)

values) is important, in order to allocate resources more effectively. As posts get more and more interactions, airline companies may want to filter the ones that really need answering, so it can get a more adequate level of customer service being demanded. In addition, having well-defined and disseminated public relation rules about which posts the airline's staff should not engage in can reallocate resources for actual customer service. The posts that got answered are represented by the nodes connected to the airline company (User-TAP). All figures denote that during the strike days,

the airline company did not address many of the questions posted on Facebook.

5.2 The semantic network

Figure 5 represents the network that encompasses posts and its concepts. By using out-degree, it was possible to determine the most frequent and relevant concepts extracted from posts, to create a semantic sub-network, generated to look for specific information used in posts. This approach allowed

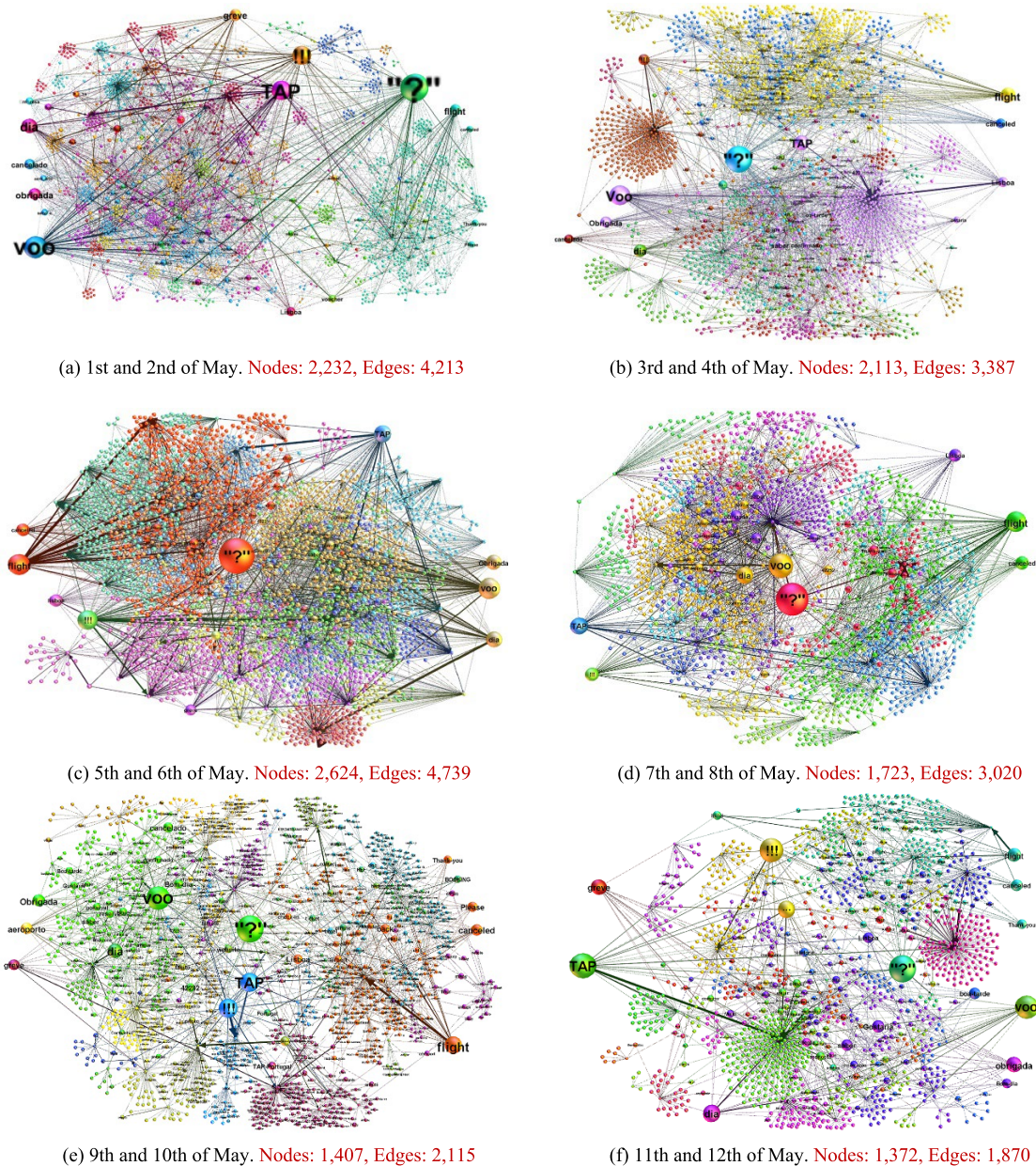


Fig. 5 Network: postconcept. **a**, **b**, **c**, **d** and **e** networks during the ten days of the strike. **f** network on the following two days. Each network is composed of nodes that represent both posts and concepts. The size of each node and label reflects the frequency of the most com-

monly used concept within the network (a more detailed visualization is available at https://drive.google.com/drive/folders/1RaH1ly1FEHuU7WtJdzfG6SwhAZgQSu9?usp=share_link)

us to identify the nodes that received the most direct connections, indicating how many times a concept had been used in the network. To this purpose, we represented the obtained results for all the 6 datasets for every two-day period.

Using a semi-automated text analysis, it was possible to extract a list of commonly used concepts and to rank them. All extracted concepts that belonged to the same post were seen as a community, by means of a modularity class algorithm. This indicated that they were more densely connected

with each other than with the rest of the network. The out-degree metric allowed to determine which nodes had the higher number of outbound links, informing on how many times a concept had been used in a post.

This approach (getting to know the network's concepts) allows airline companies, in general, to grasp which are the major issues that are being reported, to find if common solutions can be provided, optimizing customer response, by making it more cost-effective and to better allocate resources

to specific problems, when possible. According to the density and number of connections in Fig. 5 (especially during the first 8 days of the strike—networks a–d), where the same concepts are being addressed over and over again, it is reasonable to assume that the airline's customer response could have been better coordinated and that the publication of the same information that was provided many times (on the airline's website, on Facebook's main page, or by using the company's billboards at the airport, for instance) could have taken a load from answering users (one by one), using the airline's Facebook account, thus reinforcing customer service level.

Although, as mentioned in Sect. 3, users tend to mix characters and symbols to communicate quicker, especially when expressing negative emotions, the discursive exchanges found at the airline's fan page, did not resemble literature characterizations and definitions. The collected messages were quite long, formal, and well-formed (like in a commercial letter), bearing few abbreviations or emojis and, despite the apparent dissatisfaction, quite polite (interestingly, it was found that the concepts “good morning”, “good afternoon” and “thank you” were widely used). Moreover, in the posts that registered the highest values for the in-degree metric, i.e., the messages that were mostly viewed or commented on in each snapshot, it was verified that they either had a high number of concepts, or a reduced number of concepts. For example, post-134, on the 3rd and 4th of May, contained a total of 942 concepts and, in the opposite direction, post-103

on the 11th and 12th of May, only had 8 concepts. The visualization of these two messages (see Fig. 6) allowed us to conclude that, despite both showing dissatisfaction, the first was a complaint and the latter intended to express indignation/outrage denigrating the airline's image.

Messages containing a higher number of concepts not only can indicate valid assistance requests, but that messages with only a few concepts may turn out to be swearing or foul language. In this case, post-134 contained information that could be used to respond to the customer assertively, since its concepts identified the time, origin, and destination of a flight. Regarding post-103, its identification would be important, even if censoring was the decision to be made. Though not tested in this case study and thus requiring future analysis, defining thresholds for the number of concepts within each post may become important for airlines to filter valid assistance requests, from merely defamatory comments that could be automatically censored or removed from the network, as well as the user that made them.

6 Final remarks

Classic decision support systems are usually seen as the major trend in decision support, but the use of OSN's and associated technology are driving the creation of new frameworks to simplify data manipulation and decision-making and support, especially when such tools are available for

Post-134 (a) 3rd and 4th of May	Post-103 (b) 11th and 12th of May
TRES TRES MECONTENTES ET EN COLERE !!! Nous avons réservé un voyage de 7 jours à Albufeira du 24-05 au 01-06 organisé par Jet Tours ... Les vols devaient être assurés par votre compagnie. A l'aller peu de problèmes les voyages Paris-Orly-Lisbonne et Lisbonne-Faro se sont bien déroulés juste une heure de retard sur le 1er vol et le transfert vers le 2e a été très court. Au retour notre navette venant de l'hôtel est arrivée à 9h le 1-05 pour notre vol prévu à 11h15 à Faro. La télévision est présente et interroge les voyageurs. Nous apprenons que TAP est en grève. Le vol est annoncé incertain. Au bout d'un certain temps, on nous annonce que le vol va partir, mais les bagages seront déchargés à Lisbonne. Nous embarquons et partons avec une heure de retard. Arrivée à 12h40. Récupération des bagages très longue. Ensuite au pas de course nous arrivons vers une hôtesse qui nous dirige vers l'enregistrement en coupe file. Malheureusement l'enregistrement est clos ! le temps que nous avons perdu à récupérer les bagages et le retard sur le vol précédent nous a fait loucher ce vol. Après avoir fait la queue à plusieurs guichets renseignements voyageurs transfert on nous dit que les prochains vols sont annulés ou en surbooking. Réédition du billet électronique sur un vol à 19h35. Nous nous présentons à l'embarquement et oh ! horreur ! vol ANNULÉ ! Puis c'est le parcours du combattant on se présente à plusieurs comptoirs pour finir au transfert ou des dizaines de personnes font la queue. L'attente est très longue. La police est sur place... pourquoi ? A notre demande, on nous propose des tickets repas à 6€ très dérisoire pour manger correctement. On nous dit vous pouvez prendre un hôtel en attendant un vol demain vous avancez les frais on vous rembourse à hauteur de 60€ le supplément ce sera pour vous. Si vous prenez un taxi ce ne sera pas remboursé non plus ! On nous dit également de récupérer nos bagages sur les tapis d'arrivée à 22h. Lorsque nous arrivons les bagages sont posés n'importe comment au sol sans surveillance Le temps de chercher et appeler des hôtels les plus proches sont complets et horriblement chers. Recherche jusqu'à minuit. Nous passons de nombreux appels et sms coût élevé de l'étranger. Nous capitulons et nous installons sur des chaises après avoir erré dans l'aéroport pour chercher de la place car malheureusement beaucoup de personnes sont dans le même cas. A 4heures notre périple recommence nous guetons les panneaux d'affichage afin de trouver un vol pour samedi 2-05 et cette fois avec nos bagages. Nous attendons l'ouverture du guichet transferts. Celui-ci nous propose le vol de 15h45. Pourquoi ce vol si tard ? on nous dit qu'il n'est pas confirmé, qu'il peut-être encore annulé et qu'il est en surbooking et de nous y présenter très très tôt. De notre côté nous nous renseignons et il s'avère que les vols précédents soit 13h55 et 2 autres vols sont bien confirmés nous apprenons même qu'il reste de la place et qu'ils sont confirmés et décolleront ... Qu'est ce que c'est que cette pagaille ? pourquoi nous faire partir si tard ? Dans le même temps nous demandons un voucher pour le petit déjeuner et un pour le repas de midi on nous répond non un seul à la fois il faudra refaire la queue à midi ! Je vous rappelle que nous étions debout depuis 6h la veille !! A aucun moment on ne nous a proposé ne serait-ce qu'une bouteille d'eau !! Aucune aide aucun effort des employés du transfert ! Ensuite nous sommes allés nous faire réenregistrer de nouveau une queue très longue afin de lâcher nos bagages. Nous avons pris notre petit déjeuner et sommes retournés au transfert pour récupérer un voucher pour le repas de midi si on peut appeler cela un repas ! L'employée a eu le culot de vérifier si on ne venait pas une 2e fois elle a feuilleté toute la souche ! Nous avons demandé à voir son superviseur qui a refusé de nous rencontrer !! Il s'est défilé ! Après le repas nous sommes restés près des panneaux d'affichage à attendre le n° de la porte pour nous présenter dans les 1ers afin d'être sûrs d'embarquer car on nous avait dit qu'il y avait du surbooking. Nous avons quand même été surpris de voir arriver 5 personnes sans carte d'embarquement sans faire la queue se présenter aux hôtesse et se voir remettre 5 places dont 4 en classe affaire et un placée à côté de nous dans l'avion ! Comment est-ce possible ????????? L'avion a enfin décollé avec 20 minutes de retard. Arrivée 20h10 le 2-05 au lieu de 17h20 le 1-05 !! Aucune excuse de la compagnie. Comportement intolérable du personnel aux différentes guichets ! Aucune aide, aucun effort de leur part ! Nous avons perdu tout le bénéfice de notre semaine de vacances ! Nous exigeons une indemnité de 400€ dans le cas de vols au sein de l'Union européenne de plus de 1500km Conformément au règlement communautaire n° 261/2004 du 11-02-2004, entré en vigueur le 17-02-2005 les passagers refusés à l'embarquement pour ce motif doivent être indemnisés. Nous exigeons également le remboursement de toutes les dépenses engagées à l'aéroport pendant cette durée voir justificatifs. En cas de non réponse ou réponse négative je m'adresserai à la Direction Générale de l'Aviation civile. Je contacterai les associations de consommateurs, mon service juridique dans tous les cas je peux vous assurer que je ferai tout pour récupérer le dédommagement et les frais. J'attends donc très vite votre retour Cordialement	Uma Vergonha! Vende-se companhia aérea feita em cacos

Fig. 6 Example of posts with the highest and lowest number of concepts

a very low cost or even for free (which was the case of the used software in this paper). While it is true that the building of interaction networks and semantic networks has been previously explored, our study offers an approach for analyzing airline-related Facebook user data by integrating all network elements, web discourse entities, process steps, and SNA techniques, to an airline's OSN, as, to best of our knowledge, no previous studies have considered such integration. Understanding an OSN context means not only that an organization can follow what their customers do, but also that they can get much closer to what they say and want. With the explosive growth of online available social data, businesses no longer can neglect it when making decisions.

SNA can provide valuable insights into the performance of an airline's Facebook account when it is considered as a call-center service. By analyzing the interactions between the Facebook account and customers, call-center managers can identify patterns, trends, and areas for improvement. Several SNA metrics, described in Table 5, such as degree, closeness, and betweenness centrality, can be used to improve the performance of an OSN call-center.

Firstly, SNA can be used to identify influential customers by using degree centrality, eigenvector, and PageRank measures. These metrics can help call-center managers to prioritize their interactions with influential customers who

have the most connections and may influence the opinions of others. Secondly, closeness centrality can be used to improve response times and enhance the overall customer experience, by identifying customers who are closest to each other and are most likely to need a quicker response. In-degree and out-degree measures can also be used to identify the most active users in the network who need additional support.

Thirdly, betweenness centrality can be used to identify which users act as bridges between other users in an OSN call-center. This metric can help call-center managers to resolve issues more effectively and use customer feedback to drive improvements in their operations. Fourthly, betweenness centrality can be used to identify users who are most effective at resolving issues and which customers are most likely to provide valuable feedback. By prioritizing interactions with these users, call-center managers can improve the resolution at first call of issues and drive improvements in their operations.

Lastly, SNA metrics such as in-degree, out-degree, PageRank, and eigenvector centrality can help call-center managers to measure the company's performance in terms of the average speed of answer and customer satisfaction score. By monitoring these metrics, call-center managers can quickly identify the most active users in the network, the need for

Table 5 Summary of metric contributions

Topic	Description	In-Degree	Out-Degree	Closeness	Betweenness	PageRank	Eigenvector
Detection of influential customers	Identifies customers (user) who play, spend, and engage with a company or brand the most	•	•			•	•
Unanswered requests	Identifies calls (posts) that are left unanswered, "abandoned" posts, in order to provide a quicker response		•	•			
Issue-solving	Identifies if a company is effective at resolving issues and which customers are most likely to provide valuable feedback			•	•	•	•
Resolution at first call	Measures calls (posts) that are resolved on the first call without the need for the customer to follow up again		•		•		
Customer satisfaction	Measures the level of satisfaction of customers with the service provided by the call center	•		•		•	•
Answering speed	Measures the average time taken for a post (call) to be answered		•		•		

additional support, and highly influential customers who have a significant impact on customer satisfaction.

The performed analysis can help airline companies, in general, to identify troubles in traditional contact channels and to prompter prevent service breakdowns. Calculating SNA metrics and using network visualization techniques allows to use OSN's as important business aids, thus enhancing their value as business information sources for decision-making, as we ratified that SNA contributes for a deeper understanding of customers' behavior and interaction (amongst themselves and with the airline companies).

By analyzing the posts during the strike days of the referred airline company, categorizing such posts would have allowed the airline company the means to address cases complaints, namely by creating/focusing task groups for helping in reservation changes, information requests (refund procedures, available hotel options, luggage issues, etc.). The identification of concepts could have also been used by the airline company to reinforce the means for providing permanent information (over service counters or online channels) about flight cancellation; rescheduling flight procedures or even to filter valid assistance requests from defamatory ones, though further research is needed on this matter. The OSN visualization by evidencing relevant conversations would have allowed the airline company to faster address clients' dissatisfaction sentiment and whether such feeling was being aggravated or mitigated with customer service actions.

Acknowledgements This work was funded by the Portuguese Foundation for Science and Technology (FCT) (projects UIDB/00308/2020 and UIDB/05037/2020).

Author contributions MF, FA and JPC have conceived the study. MF has performed the analysis. MF, FA and JPC have written the manuscript. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

- Aggrawal N, Anand A (2022) Social networks: modeling and analysis. Taylor & Francis Limited, Boca Raton
- Alam F, Ofli F, Imran M (2020) Descriptive and visual summaries of disaster events using artificial intelligence techniques: case studies of Hurricanes Harvey, Irma, and Maria. *Behav Inf Technol* 39(3):288–318. <https://doi.org/10.1080/0144929X.2019.1610908>
- Al-deen HSS, Zeng Z, Al-sabri R, Hekmat A (2021) An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism. *Appl Syst Innov* 4(4):85. <https://doi.org/10.3390/asi4040085>
- Antunes F, Costa JP (2011) Decision Support Social Network Conference in Information Systems and Technologies (CISTI). Chaves, Portugal
- Antunes F, Freire M, Costa JP (2014) Semantic Web Tools and Decision-making. In: Zaraté P, Kersten GE, Hernández JE (eds) Group decision and negotiation: a process-oriented view, lecture notes in business information processing (LNBIP). Springer, Berlin, pp 270–277
- Antunes F, Freire M, Costa JP (2016) Semantic web and decision support systems. *J Decis Syst* 25(1):79–93. <https://doi.org/10.1080/12460125.2015.1087293>
- Arafeh M, Ceravolo P, Mourad A, Damiani E, Emanuele B (2021) Ontology based recommender system using social network data. *Futur Gener Comput Syst* 115(2021):769–779. <https://doi.org/10.1016/j.future.2020.09.030>
- Barão A (2014) The Gephi Network Splitter 3D Layout Retrieved 23 July 2022. <https://www.relationalcapitalvalue.com/gephiplugins.html>
- Bartos M, Badanik B (2019) Flying social media course. *Transp Res Procedia* 43(2019):119–128. <https://doi.org/10.1016/j.trpro.2019.12.026>
- Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. In: Proceedings of the international AAAI conference on web and social media. pp 361–362. <https://doi.org/10.1609/icwsm.v3i1.13937>
- Bechberger D, Perryman J (2020) Graph databases in action. Manning Publications, Shelter Island
- Borgatti SP (2009) 2-mode concepts in social network analysis. In: Meyers RA (ed) Encyclopedia of complexity and system science. Springer, Larkspur, pp 8279–8291
- Burguillo JC (2018) Self-organizing coalitions for managing complexity: agent-based simulation of evolutionary game theory models using dynamic social networks for interdisciplinary applications. Springer, Cham. <https://doi.org/10.1007/978-3-319-69898-4>
- Camacho D, Panizo-Lledot Á, Bello-Ortiz G, González-Pardo A, Cambria E (2020) The four dimensions of social network analysis: an overview of research methods, applications, and software tools. *Inf Fusion* 63:88–120. <https://doi.org/10.1016/j.inffus.2020.05.009>
- Dutta Das D, Sharma S, Natani S, Khare N, Singh B (2017) Sentimental analysis for Airline Twitter data. In: IOP conference series: materials science and engineering, 263, 042067. <https://doi.org/10.1088/1757-899X/263/4/042067>
- Fani H, Jiang E, Bagheri E, Al-Obeidat F, Du W, Kargar M (2020) User community detection via embedding of social network structure and temporal content. *Inf Process Manag* 57(2):102056. <https://doi.org/10.1016/j.ipm.2019.102056>
- Fedushko S, Molodetska K, Syerov Y (2023) Decision-making approaches in the antagonistic digital communication of the online communities users. *Soc Netw Anal Min* 13(1):18. <https://doi.org/10.1007/s13278-022-01021-4>
- Freire M, Antunes F, Costa JP (2017) A semantics extraction framework for decision support in context-specific social web networks. In: Linden I, Liu S, Colot C (eds) Decision support systems VII data, information and knowledge visualization in decision support systems. Springer, Cham, pp 133–147. https://doi.org/10.1007/978-3-319-57487-5_10
- Freire M, Antunes F, Costa JP (2022) Getting decision support from context-specific online social networks: a case study. *Soc Netw Anal Min* 12(1):41. <https://doi.org/10.1007/s13278-022-00870-3>
- Freire M, Antunes F, Costa JP (2015) Exploring social network analysis techniques on decision support. In: Mesquita A, Peres P (eds) ECSM 2015 2nd European conference on social media. Academic Conferences and Publishing International Limited, pp 165–173

- Fu X, Luo J-D, Boos M (2017) Social network analysis: interdisciplinary approaches and case studies. Taylor & Francis Group, Boca Raton
- Galety MG, Atroshi CA, Balabantaray B, Mohanty SN (2022) Social network analysis: theory and applications. Wiley, Hoboken
- Giarelis N, Nakakis N, Karacapilidis N (2021) A comparative assessment of state-of-the-art methods for multilingual unsupervised keyphrase extraction. In: Maglogiannis I, Macintyre J, Iliadis L (eds) Artificial Intelligence Applications and Innovations. Springer, Cham, pp 635–645. https://doi.org/10.1007/978-3-030-79150-6_50
- Goritz A, Kolleck N, Jörgens H (2019) Education for sustainable development and climate change education: the potential of social network analysis based on Twitter data. Sustainability 11(19):5499. <https://doi.org/10.3390/su11195499>
- Hansen DL, Shneiderman B, Smith M, Himelboim I (2020) Analysing social media networks with NodeXL: insights from a connected world. Morgan Kaufmann, Burlington. <https://doi.org/10.1016/B978-0-12-817756-3.09995-7>
- Herring S (2013) Discourse in Web 2.0: familiar, reconfigured, and emergent. In: Tannen D, Trester A-M (eds) Discourse 2.0: language and new media. Georgetown University Press, Washington, pp 1–25
- Isson JP (2018) Unstructured data analytics: how to improve customer acquisition, customer retention, and fraud detection and prevention. Wiley, Hoboken
- Kok S, Rogers R (2017) Rethinking migration in the digital age-translocalization and the Somali diaspora. Global Netw 17(1):23–46. <https://doi.org/10.1111/glob.12127>
- Kumari MV, Prajna B (2021) Collaborative classification approach for airline tweets using sentiment analysis. Turk J Comput Math Educ 12(3):3597–3603. <https://doi.org/10.17762/turcomat.v12i3.1639>
- Liu X, Min Q, Wu D, Liu Z (2020) How does social network diversity affect users' lurking intention toward social network services? Role Perspect Inf Manag 57(7):1–16. <https://doi.org/10.1016/j.im.2019.103258>
- Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality. In: Burges CJC, Bottou L, Welling M, Ghahramani, Weinberger KQ (eds) Proceedings of the 26th international conference on neural information processing systems. Press Associates Inc., New York, pp 3111–3119. <https://doi.org/10.5555/2999792.2999959>
- Moser C, Groenewegen P, Huysman M (2013) Extending social network analysis with discourse analysis: combining relational with interpretive data. In: Ozyer T, Rokne J, Wagner G, Reuser A (eds) The influence of technology on social network analysis and mining. Springer, New York, pp 547–561. https://doi.org/10.1007/978-3-7091-1346-2_24
- Naseem U, Khan SK, Razzak I, Hameed IA (2019) Hybrid words representation for airlines sentiment analysis. In: Liu J, Bailey J (eds) AI 2019 advances in artificial intelligence—32nd Australasian joint conference. Springer, Adelaide
- Pennington D (2017) Coding of non-text data. In: Sloan L, Quan-Haase A (eds) The Sage handbook of social media research methods. SAGE, London, pp 232–250
- Pennington J, Socher R, Manning C (2014) GloVe: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). Association for computational linguistics, Doha, Qatar, pp 1532–1543. <https://doi.org/10.3115/v1/D14-1162>
- Piedrahita P, Borge-Holthoefer J, Moreno Y, González-Bailón S (2017) The contagion effects of repeated activation in social networks. Elsevier: Social Networks, Amsterdam. <https://doi.org/10.1016/j.socnet.2017.11.001>
- Rane A, Kumar A (2018) Sentiment classification system of Twitter Data for US Airline Service Analysis. In: 2018 IEEE 42nd annual computer software and applications conference (COMPSAC). IEEE, pp 769–773. <https://doi.org/10.1109/compsac.2018.00114>
- Savic M, Ivanovic M, Jain LC (2019) Complex networks in software, knowledge, and social systems. Springer, Cham. <https://doi.org/10.1007/978-3-319-91196-0>
- SimpliFlying. (2019). Airline Social Media Outlook Report 2019. SimpliFlying
- St-Onge J, Renaud-Desjardins L, Mongeau P, Saint-Charles J (2022) Socio-semantic networks as mutualistic networks. Sci Rep 12:1889. <https://doi.org/10.1038/s41598-022-05743-5>
- Tiwari P, Yadav P, Kumar S, Mishra BK, Nguyen GN, Gochhayat SP, Singh J, Prasad M (2019) Sentiment analysis for airlines services based on Twitter dataset. In: Nilanjan D, Borah S, Babo R, Ashour AS (eds) Social network analytics. Academic Press, Cambridge, pp 149–162. <https://doi.org/10.1016/B978-0-12-815458-8.00008-6>
- Wang L, Ren X, Wan H, Yan J (2020) Managerial responses to online reviews under budget constraints: whom to target and how. Inf Manag 57(8):1–13. <https://doi.org/10.1016/j.im.2020.103382>
- Wankhade M, Rao ACS, Kulkarni C (2022) A survey on sentiment analysis methods, applications, and challenges. Artif Intell Rev 55:5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>
- Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511815478>
- Wu Y, Lu R, Su H, Shi P, Wu Z-G (2017) Synchronization control for large-scale network systems. Springer, Cham. <https://doi.org/10.1007/978-3-319-45150-3>
- Yang J, Xiu P, Sun L, Ying L, Muthu B (2022) Social media data analytics for business decision making system to competitive analysis. Inf Process Manag 59(1):15. <https://doi.org/10.1016/j.ipm.2021.102751>
- Yuksel AS, Tan FG (2018) A real-time social network-based knowledge discovery system for decision making. Automatika 59:261–273. <https://doi.org/10.1080/00051144.2018.1531214>
- Zhao H, Liu Z, Yao X, Yang Q (2021) A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. Inf Process Manag 58(5):14. <https://doi.org/10.1016/j.ipm.2021.102656>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.