Abstract: The recent emergence of e-government and online social media offers opportunities for governments to meet the demands and expectations of citizens, to provide value-added services and overcome barriers of reduced public budgets. Twitter is the most popular microblogging platform that can facilitate interaction and engagement. It is widely used by government agencies, public affairs practitioners, non-government organizations, members of Parliament and politicians. The paper aims to explore the use of Twitter by government agencies in Greece and record Twitter followers’ preferences regarding which accounts they follow. The paper records 27 Greek e-government Twitter accounts and their 107,107 followers. It uses a data mining technique, association rules and two multivariate statistical methods, multidimensional scaling and cluster analysis and proposes the use of a similarity measure, suitable for describing Twitter account proximity. In this way, the paper locates accounts that share followers. Groups of Twitter accounts are located, and their common orientation is described. The analysis not only describes Twitter account similarities and group formation, but to some extent, the followers’ preferences and habits of obtaining information through Twitter, as well.

Keywords: e-government; Twitter; Greece; followers; association rules; multidimensional scaling; cluster analysis; groups of accounts; shared followers
1. Introduction

Information and communication technologies have changed the process of governing in the world. Governments worldwide have responded to the challenging opportunities offered by e-government in order to form an administration that engages with citizens and consults them, to meet the increasing demands and expectations of citizens, to overcome the major barriers of reduced public budgets and to provide value-added services to their citizens [1–3]. An open government is committed to “bridging the gap” between the government and its citizens [2] and to the principles of “transparency, public participation, and collaboration” [4].

Nowadays, governments have started embodying social media technologies and applications in their e-government strategies in an attempt to maximize the Web 2.0 benefits and to keep up with current trends [5]. The use of social media by public agencies may enhance participation, transparency, accountability, integration, communication and collaboration, promote civic engagement, revitalize dialogue between citizens and government and speed up the pace of innovation [6–8].

Twitter is a microblogging service that allows users to share information via short messages with a maximum of 140 characters in length and to answer the questions: “What’s happening?”, previously, and “What are you doing now?” [9, 10]. The restricted size of the text implies that tweets may not necessarily contain well-formed ideas or developed context, but they are complete enough to be coherent [11]. Moreover, the quick nature of Twitter results in a constantly updated stream of short messages ranging from breaking news, real-life events, like revolutions and natural disasters, global discussions, shared links and thoughts on life. People use Twitter to ask for directions, support, advice, to communicate, to ask questions and to validate ideas by discussing them with others. Thus, Twitter allows for large quantities of concise and diverse concepts to evolve [11–14].

Twitter affords a variety of functions; users can set their updates to private, or they can allow the entire cyberspace area of Twitter, the “Twitterverse”, to view their pages [15]. The vast majority of Twitter accounts are public [16]. Users may also choose to follow other users; thus, the structure of followers forms a large network [17]. However, the act of “following” users on Twitter is not mutual. An “asymmetric” model of human relationships exists [18]. In this vein, Twitter does not properly exhibit a “traditional” social network structure [14]. In the Twitter vocabulary, to write a tweet addressing a specific user is called a mention, @reply is a tweet directed at a certain user, RT stands for retweet and a “#” followed by a word represents a hashtag. Retweets allow users to rebroadcast content generated by others, thereby raising the content’s visibility [19], while hashtags are grouping tweets by topic, allowing users to annotate tweets with metadata [20]. According to Sousa et al. [10], users are connected in an implicit or explicit manner via retweets, #hashtags, @replies and follower relationships. Moreover, third parties offer services that add functionality to the Twitter site, such as the ability to donate money or rank the influence of its users [21].

Twitter is one of the most used social media applications by government agencies, public affairs practitioners and politicians [22]. Regarding e-government, Greece lags behind other EU nations [23] and also in global terms. The United Nations’ e-government survey [24] ranked Greece’s e-government project as number 37 worldwide; Greece increased its world ranking from the 41st position in 2010. Greece is now included in the list of countries that provide a statement “follow us on Facebook or Twitter” at government websites [24].
The paper aims to investigate how followers distribute among e-government accounts and to record citizens’ preferences regarding which Twitter accounts they follow. The paper, besides reporting descriptive statistics regarding followers, proposes a methodology for locating and describing groups of accounts which share followers.

The paper is organized as follows:

- Section 2 reports on the related work on the use of social media and especially Twitter in e-government;
- Section 3 presents the recorded data and a brief reference to the analysis that follows;
- Section 4 provides descriptive statistics of the data and it introduces the use of an association rules approach as a complementary descriptive technique for large data sets. Next, a statistical approach is introduced in order to describe patterns of accounts that share followers. Findings are presented and discussed. The method aims to provide groupings of accounts by measuring similarities between accounts regarding the followers they share. By locating such groups of Twitter accounts, it may be easier to understand what the citizens’ habits are regarding how they choose to follow Twitter accounts and use the particular information sources.

2. Related Work

2.1. e-Government and Social Media

e-Government can be defined as “the use of information and communications technologies (ICTs), and particularly, the Internet, as a tool to achieve better government” [25]. The appropriate use of ICTs plays a crucial role, as the real benefit of e-government does not lie in the use of technology per se, but in its application in order to transform processes and enhance public sector information delivery [26]. e-Government services may be provided through the Internet, kiosks, integrated service centers and mobile devices [27] and different platforms, like websites, social media and mobile applications, in order to maximize citizens’ reach.

Social media are now employed by the majority of Internet users [28]. Facebook has more than 1.1 billion monthly active users and 751 million monthly mobile users. Google+ is now the second largest social network, with 359 million monthly active users [29]. Twitter has 554,750,000 active registered users and is the fastest growing social network by active users. One of Twitter’s faster growing age groups is 55 to 64 years old with a 79% growth rate since 2012 [30]. Smith [31] mentioned that one third of all online adults in the USA used social media to keep informed about government activities. In this vein, Pirolli et al. [32] claimed that as the number of social media users increases, there has been a growing interest by governments in using them.

Social media have been recognized as an important information source by many governments [33] that provide them with the ability to acquire and distribute information very quickly [34]. Chun and Luna-Reyes [35] provided a review of basic Web 2.0 technologies, categorizations and contributions. According to them “social media use in government is considered a technological innovation and a transforming agent in generating citizen engagement from campaigns and grassroots-activism to shared governance promoting democracy” ([35], p. 42). Boyd [36] uses the term social media as an “umbrella term that refers to the set of tools, services, and applications that allow people to interact...
with others using network technologies” ([36], p. 92). According to this definition, communication is arguably the most important factor, while dissemination of news forms is only a small part [37]. With the use of social media, a “virtual public sphere” [38] appears, offering a platform for political deliberation and participation [37]. Social media have the ability to encourage increased engagement with citizens [39], especially with underrepresented segments of the population that previously were difficult to reach [28], and facilitate the democratic value of free speech and participation for populations under authoritarian rule [40].

However, the concept of political participation is multidimensional and embodies actions, such as voting, citizens’ involvement in election campaigns, attending meetings or attempting to access information on different political parties, engagement in grassroots politics within their local communities, attending civil protests, signing petitions, joining interest groups that engage in lobbying or political advocacy [41]. Social media possess the potential to allow users to get involved in social-political processes, to air their views, to express anger and dissatisfaction, to share their own stories and concerns, to become important catalysts of collective action processes and to activate their own social networks [41,42]. Governments may actively use social media in order to allow citizens to interact with organizations [43] and to enhance open corporate dialog [44]. The public may participate in policymaking and provide governments with the collective knowledge, ideas and expertise of the population [45]. In this vein, users may be consulted for certain projects, decisions or activities via forums, polls and other social media [43], participate directly in the process of communication through the contribution of content by commenting or tagging [46] and provide feedback [47]. Thus, government’s effectiveness should be enhanced and the quality of its decisions improved [45].

2.2. Twitter in e-Government

In the e-government context, Twitter is increasingly adopted by government departments, Members of Parliament, millions of businesses, non-governmental organizations and individuals [48]. Government agencies using Twitter find it “an effective, efficient, timely and valuable tool to get the word out” claimed Wigand ([49], p. 66). An especially promising aspect is that individuals and politicians interact within the same symbolic framework using messages of identical length, the same type of language, the same external references and, frequently, the same content use of hashtags and retweets [50]. However, little research effort has been devoted to investigating government use of Twitter [51].

One of the first studies regarding the use of Twitter by governments was that of Golbeck et al. [52], who were interested in the type of content that legislators, particularly members of the United States Congress, were posting. They analyzed the content of over 6000 posts from all members of Congress. Legislators use Twitter as a vehicle for self-promotion and do not provide new insights into government or the legislative process or improve transparency. However, according to the findings, Twitter was facilitating direct communication between Congress people and citizens.

Heverin and Zach [53] investigated the use of Twitter by city police departments in U.S. cities with populations greater than 300,000. They analyzed 4,915 posts and found that police departments primarily use Twitter to share information about their departments, events, traffic, safety awareness and crime prevention and primarily crime and incident related information. Use of Twitter by police in the U.K. was examined by Crump [54] who aimed at investigating the structure of networks formed
and the content of the messages. The article concludes that exchanges within networks are infrequent, and the nature of Twitter means that conversations are difficult to join. However, Twitter has been used cautiously and as a reinforcement for existing means of communication.

Alam and Lucas [51] examined Australian government use of Twitter by analyzing six government agency tweets. The analysis showed that agencies were primarily using Twitter to disseminate information, particularly links to news articles about themselves, and to report on their activities.

Waters and Williams [22] by using the models of public relations as a framework examined how government agencies are using Twitter to communicate with their audiences. A total of 1800 tweets from 60 government agencies were analyzed using content analysis. Findings indicate that government agencies sought to inform and educate their followers, while interactivity on Twitter relied on one-way communication rather than two-way symmetrical conversations.

Small [55] performed an analysis of the Canadian Government’s use of Twitter. The Canadian government is using Twitter as part of its well-established e-government strategy. Based on content analysis, she claimed that service delivery characterizes government tweets.

An analysis of the Twitter activity of the Ministry for Food, Agriculture, Forestry and Fisheries (MFAFF), one of Korea’s government organizations, was performed by Cho and Park [56] by using content analysis, semantic network analysis and ego network analysis. Their findings suggested that Twitter can function as an effective information delivery channel for government agencies and a communication arena where socialization-oriented conversations can be mutually exchanged. However, as the study concerns only one governmental organization, the generalization of the findings is limited.

Unsworth and Townes [57] reported on the use of Twitter by the United States Department of Agriculture (USDA). They applied a mixed-methods approach based on grounded theory, social network analysis, content analysis and discourse analysis. The findings suggest that there is no clear evidence of discourse, in the sense of exchanging tweets between the USDA and other users.

Panagiotopoulos and Sams [58] studied Twitter accounts maintained by 187 U.K. local government authorities. The analysis was conducted by examination of the Twitter networks developed by the accounts, and it was followed by a structural analysis of more than 296,000 tweets. Regarding the Twitter network, findings indicate that it gives U.K. local government access to a diverse group of stakeholders beyond networked citizens at the local level; while regarding the interactivity with citizens, Twitter accounts attempt to satisfy quite complicated information needs. As far as the content produced by the accounts is concerned, the content is localized and temporal. In addition to this, agencies are tweeting about local democratic processes, possibly in an attempt to inform citizens about them or even stimulate their engagement.

3. Data Collection and Description

This is exploratory research, aiming to locate Twitter accounts that share followers, and in this way, to conclude on how the citizens choose what accounts to follow. Furthermore, the paper proposes two quantitative techniques: association rules are used as a complementary descriptive technique, which can help us describe the large picture, and multidimensional scaling followed by hierarchical cluster analysis are used to locate combinations of accounts sharing followers. The questions that this study aims to answer are:
1. How many citizens follow each account?
2. How many of them follow more than one account?
3. How many accounts share followers?
4. What are the combinations of accounts that are followed by the same followers?
5. Do followers choose which accounts to follow by means of the accounts’ thematic similarity?

During 10–11 May 2013, the followers of 27 Greek e-government Twitter accounts that have a national appeal and interest were recorded. Recording of the accounts was based on the official registers and agency records of the Greek government. The accounts referred to the general government, the ministries with its services and their secretariats, along with the operational programs. Two independent authorities and one official forum that deal with e-government matters were included. To find the accounts, as the first step, the official web sites of the organizations were used, in order to search for links to Twitter. Next, we used Google search with specific key words. As the last step, we used Twitter’s own search engine. The recorded accounts are described in Table 1. Table 1 also describes the intended purpose and the target audience of each account (G2C: Gov-to-Citizen, G2B: Gov-to-Business, G2G: Gov-to-Gov), in order to enrich the finding of the analysis. The intended purpose of the accounts drew upon a list of main e-participation areas. Table 1 also reports the number of followers of each account along with the percentage of them that also follow other accounts.

Table 1. The 27 e-government Twitter accounts included in the analysis.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Intended purpose (target audience)</th>
<th>Followers</th>
<th>Percentage of followers that also follow other accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCoastGuard</td>
<td>Official account of the Greek coastguard.</td>
<td>Information Provision (G2C)</td>
<td>1,522</td>
<td>88%</td>
</tr>
<tr>
<td>Diavgeia</td>
<td>Account of the Greek Ministry of Administrative Reform and e-government, which aims to achieve maximum publicity of government policy (Ministry of Interior).</td>
<td>Information Provision, Consultation (G2C)</td>
<td>2,256</td>
<td>91%</td>
</tr>
<tr>
<td>Digitalgreece</td>
<td>Digital Greece 2020 Forum.</td>
<td>Community Building/Collaborative Environments, Information Provision, Consultation (G2C, G2B)</td>
<td>469</td>
<td>77%</td>
</tr>
<tr>
<td>DigiPlanGR</td>
<td>Secretariat for Digital Planning.</td>
<td>Information Provision (G2C, G2B)</td>
<td>738</td>
<td>91%</td>
</tr>
<tr>
<td>e_ekpaideusi</td>
<td>Operational Programme for Education and Lifelong Learning from the Greek Ministry of Education.</td>
<td>Information Provision (G2C, G2B)</td>
<td>77</td>
<td>78%</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Intended purpose (target audience)</td>
<td>Followers</td>
<td>Percentage of followers that also follow other accounts</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>Madcgr</td>
<td>Operational Programme from General Secretariat for Telecommunications and Postal services.</td>
<td>Information Provision (G2C, G2B)</td>
<td>24</td>
<td>83%</td>
</tr>
<tr>
<td>Gsaeduu</td>
<td>Greek General Secretariat of Lifelong Learning (Ministry of Education).</td>
<td>Information Provision (G2C, G2B)</td>
<td>20</td>
<td>60%</td>
</tr>
<tr>
<td>GGMEofficial</td>
<td>Official Twitter account of the General Secretariat of the Media.</td>
<td>Information Provision, References (G2C, G2B)</td>
<td>76</td>
<td>79%</td>
</tr>
<tr>
<td>GGNeasGenias</td>
<td>General Secretariat for youth.</td>
<td>Information Provision (G2C)</td>
<td>2,392</td>
<td>87%</td>
</tr>
<tr>
<td>PressParliament</td>
<td>Press and Parliamentary Reporting from Greek Parliament.</td>
<td>Information Provision (G2C)</td>
<td>508</td>
<td>89%</td>
</tr>
<tr>
<td>Govgr</td>
<td>The official account of the Greek Government website.</td>
<td>Information Provision (G2C)</td>
<td>15,931</td>
<td>89%</td>
</tr>
<tr>
<td>Hellenicpolice</td>
<td>Official account of Hellenic police.</td>
<td>Information Provision (G2C)</td>
<td>31,423</td>
<td>69%</td>
</tr>
<tr>
<td>HellasYAP</td>
<td>Service that aims to implement the government policy for the introduction, implementation and development of information and technology in the public sector (Greek Ministry of Administrative Reform and e-government).</td>
<td>Information Provision (G2G)</td>
<td>9</td>
<td>100%</td>
</tr>
<tr>
<td>MinDevGR</td>
<td>Official account of the Greek Ministry of Development.</td>
<td>Information Provision (G2C)</td>
<td>5,312</td>
<td>95%</td>
</tr>
<tr>
<td>GreeceMFA</td>
<td>Official account of the Greek Ministry of Foreign Affairs.</td>
<td>Information Provision (G2C)</td>
<td>19,586</td>
<td>77%</td>
</tr>
<tr>
<td>Yyka_gov</td>
<td>Official account of the Greek Ministry of Health.</td>
<td>Information Provision (G2C)</td>
<td>1,681</td>
<td>82%</td>
</tr>
<tr>
<td>Minfingr</td>
<td>Official account of the Greek Ministry of Finance.</td>
<td>Information Provision (G2C)</td>
<td>11</td>
<td>73%</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Intended purpose (target audience)</td>
<td>Followers</td>
<td>Percentage of followers that also follow other accounts</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>-----------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>Yp_Paideias</td>
<td>Account with RSS feeds from the Greek Ministry of Education.</td>
<td>Information Provision (G2C)</td>
<td>1,785</td>
<td>76%</td>
</tr>
<tr>
<td>YPEKA</td>
<td>Official account of the Greek Ministry of Environment.</td>
<td>Information Provision (G2C)</td>
<td>3,362</td>
<td>82%</td>
</tr>
<tr>
<td>OpenGovGr</td>
<td>Official account of the Greek Ministry of Administrative Reform and e-government.</td>
<td>Information Provision, Consultation (G2C)</td>
<td>8,573</td>
<td>92%</td>
</tr>
<tr>
<td>Prime ministerGR</td>
<td>Account of the Prime Minister of the Hellenic Republic.</td>
<td>Information Provision (G2C)</td>
<td>56,804</td>
<td>45%</td>
</tr>
<tr>
<td>Pyrosvestiki</td>
<td>Official Twitter account of the Fire Brigade.</td>
<td>Information Provision (G2C)</td>
<td>15,636</td>
<td>83%</td>
</tr>
<tr>
<td>Synigoros</td>
<td>Mediate between citizens and public services for the protection of civil rights and the fight against maladministration (independent authority).</td>
<td>Information Provision, Mediation (G2C)</td>
<td>1,802</td>
<td>75%</td>
</tr>
<tr>
<td>StartupGreece</td>
<td>Online information and networking space, aiming to help create a new generation of entrepreneurs in Greece (Ministry of Development).</td>
<td>Information Provision, Collaborative Environments (G2B)</td>
<td>6,035</td>
<td>63%</td>
</tr>
<tr>
<td>VisitGreecegr</td>
<td>The official Twitter account of the Greek National Tourism Organization.</td>
<td>Information Provision (G2C, G2B)</td>
<td>22,707</td>
<td>40%</td>
</tr>
<tr>
<td>Synigkatanaloti</td>
<td>Account of an independent authority, which assumes inexpensive mediation to amicably resolve consumer disputes with suppliers of goods or services.</td>
<td>Information Provision, Mediation (G2C)</td>
<td>38</td>
<td>71%</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td></td>
<td></td>
<td>1,785</td>
<td>79%</td>
</tr>
<tr>
<td><strong>Average (SD)</strong></td>
<td></td>
<td></td>
<td>7,389</td>
<td>78% (14%)</td>
</tr>
</tbody>
</table>

It should be noted that the account or the Prime Minister seems to be a personal account, and one could argue that it should not be included in the analysis. However, a link to this account along with links to the accounts of Ministries and Secretariats is provided on the official site of the Greek government. We think that because of this and because Twitter use in Greece is at an early stage, it is
meaningful to include the Prime Minister’s account. By doing so, we include as many accounts as possible and present the citizens’ choices within the whole spectrum of available sources. In any case, this paper introduces methodological approaches that could be used either by including the specific account or not.

The data consists of a 107,107 by 27 binary input matrix, where one is placed in a cell if a citizen follows the specific account. In this way, a citizen can be represented to follow as many accounts as there are columns in the matrix; in this case, unities are placed in the relative cells. This matrix will be analyzed both by using descriptive statistics and association rules, and further analysis will be done using multidimensional scaling and hierarchical cluster analysis.

4. Analysis and Findings
4.1. Descriptive Statistics

Figure 1 presents the number of followers for each account (see also Table 1). The numbers range from nine to 56,808. The Prime Minister’s account has the most followers. Next, the police account is followed by 31,423 and, interestingly, in the third place, the VisitGreecegr account is followed by 22,707 followers. The Ministry of Foreign Affairs account, the Fire Brigade account, and open government accounts come next.

Figure 1. Followers of the Greek e-government Twitter accounts.

This section also reports on the number of followers with regards to the number of e-government accounts they follow. The distribution is reported in Figure 2. Figure 2 presents the medians of the percentages of followers that follow just one account, two accounts, three accounts, etc. The medians are computed after calculating for each account how many of its followers also follow other accounts.
We chose to use the median and not the total percentage for the total pool of the followers, because there is a great diversity, dispersion and skewness of such percentages among the accounts. Therefore, accounts with more followers could affect the total percentage, and if these accounts have a much different distribution of followers, the whole picture would be biased. For example, the Prime Minister’s account has 56,804 followers, and 55% of them do not follow any other account; see also Table 1. The VisitGreece account has 22,707 followers, and 60% of them do not follow any other account (Table 1). On the other hand, these percentages range from zero to 60%. Therefore, a total percentage or a mean would not represent the data accurately.

Figure 2 presents the median percentages of the followers that follow one, two or more accounts. The median percentage of the followers that follow only one account is 22%. The median percentage for those who follow three accounts is 11%; for those who follow three accounts, it is 9.3%. Next, medians drop down more smoothly as the number of accounts increase. The last column of Table 1 presents what percentage of the followers of each account also follows other accounts. The average and the median percentage is nearly 80%. It is very common for accounts to share followers. The Prime Minister’s account and VisitGreece account are the two accounts that share relatively fewer followers that the other accounts.

Regarding the number of citizens, who follow just one account, 45% of them follow the Prime Minister’s account, 20% follow the VisitGreece account, 14% follow the Police account, 6.3% follow the Ministry of Foreign Affairs account, 4% follow the Fire Brigade account, 3.2% follow the Startup Greece account, 2.6% follow the account of Greek government and 1% follow the open government account. Percentages are smaller than 1% for the rest of the accounts.

Figure 2. Distribution of the number of accounts that share followers, using medians.

4.2. An Association Rules Approach

This subsection continues the analysis of the previous one. The proposed procedure uses association rules to describe the combinations of accounts followed by the same followers. Association rules are a data mining technique, which are used to manage large data sets and allow the counting of how many
followers follow combinations of accounts and, eventually, arriving at the best rules, such as “those who follow this set of accounts, also follow this account”. At an elementary stage, the procedure produces results of the form: followers of A are also followers of B, and followers of C, etc. For example, the procedure produces combinations of the form, aktofilaki, Helenic.police, pirosvestiki (974), to describe that 974 followers are common to the accounts, Police, Coastguard and Fire Brigade. It produces an extended list of combinations of Twitter accounts, which are followed by the same followers. It is this kind of finding that we are interested in to record and describe all possible combinations of account sharing followers.

To proceed with the technique, we have to define two indexes: support and confidence. Support is the minimum percentage of followers that an account or combination of accounts should have in order to be listed. For example, if support equals 0.50, then any combination of accounts should have at least half of the total number of followers of the data set. We set support equal to 0.00019, in order to allow accounts or a combination of accounts to be included, even if they have as many as 20 followers. For two combinations of accounts, confidence is the percentage of followers of the first combination that are also followers of the second. We set this percentage to 10%, so confidence equals 0.10. Therefore, we considered two combinations of accounts to share followers, even if 10% (or more) of the followers of the one are also followers of the other. Of course confidence can vary. We chose such a low value to allow for as many cases to be recorded as possible. In the abovementioned example, aktofilaki, Helenic.police, pirosvestiki (974), support is 974/107,107. This equals almost 1%. Confidence is equal to the proportion of the followers of the two accounts, who are also followers of the third.

We proceeded with generating association rules for the 107,107 followers of the 27 Greek e-government Twitter accounts, using WEKA (Waikato Environment for Knowledge Analysis) data mining software developed and distributed by the University of Waikato, New Zealand [59]. We used association rules in this context as a sophisticated, yet descriptive, technique to better present shared followers.

The procedure produced combinations of accounts consisting of just one account, two accounts, etc., ending up with combinations of thirteen accounts. Figure 3 presents the numbers of combinations with regards to the number of members (Twitter accounts) in each combination. For example, 25 accounts constitute 25 combinations of one member. There exist 203 combinations of two accounts that share followers. These 203 combinations are of the form “Helenic.police, pirosvestiki”. They are combinations containing two accounts each, which have the properties of support and confidence. That is, their common followers are at least 20%, and at least 10% of the followers of the first account are also followers of the second account. Further, there are 1074 combinations of three accounts that share followers, etc. The most populous combination category is that of seven accounts that share followers. There are 16,746 combinations of 13 accounts that have at least 20 followers in common (support). The distribution presented in Figure 3 is nearly symmetric, tending to by normal. It generally describes the procedure and the results, but of course, it does not provide details on what the combinations are. At this stage, the whole picture is captured only by describing the combination distributions. Figure 3 describes the ability of the accounts to share followers and form combinations with other accounts that share followers with them. In Figure 3, if a combination of 13 accounts is included in the rightmost bar, then part of this combination (having 12 accounts) is also included and presented in the second
last bar. The left bars describe combinations with fewer members than those described by bars on the right and probably some additional new combinations. For a combination of k member accounts, all the combinations with fewer member accounts are also included and presented in the graph. It is interesting that when moving from combinations of 13 members to combinations of seven members, the number of combinations increases: 15–185––16,746, while the number decreases when moving from combinations of seven members to combinations of one member. The mode of this distribution is seven, which means that it is very likely to find combinations of seven accounts that share followers. The mean equals seven and the standard deviation equals 1.7. Therefore, on average, there will be combinations of seven accounts, and also, combinations ranging from five to nine accounts are frequent. By raising support (and confidence), we can limit both the number and size of combinations and arrive at combinations of accounts that share more followers.

Figure 3. Distribution of combinations of accounts that share followers.

Figure 4 presents the average number of followers for every combination size. The graph presents average numbers of followers counted for combinations of accounts at each step of the procedure. Followers that follow small combinations can also follow combinations of more accounts. Combinations of five to nine accounts that share followers have an average number of shared followers that ranges from 94 to 35. Only combinations of up to four accounts have an average number of shared followers that is greater than 100, and only combinations of two accounts have an average of more than 1000 shared followers.

In Figure 5, we present eight box-plots. Box-plots are constructed using minimum and maximum values and the median, M, 25% percentile (Q1) and the 75% percentile (Q3). To construct the box-plots, we considered all the observations reported in Figure 3, that is a total of 72,586 cases. Additionally, we considered for every Twitter account how many times it belongs to a combination of two, three, four, and so on, accounts. Then, we constructed 27 box-plots, one for each account. Each box-plot represents the distribution of the number of combinations to which a specific account belongs. Because many such box-plots are the same for several accounts, for economy, we grouped together accounts having the same box-plots. Eight distribution categories are formed. In Figure 5, the eight
categories are presented along with their associated box-plots. Box-plots are presented in descending order: from distributions of accounts that can have members in combinations of up to 13 members, to distributions of accounts that have members in combinations of fewer members. A description of the distribution categories is presented in Table 2.

**Figure 4.** Average number of followers for every combination size.

**Figure 5.** Box-plots of Twitter accounts regarding their membership in combinations of accounts sharing followers.
Table 2. The eight distribution categories of e-government Twitter accounts according to the number of combinations to which the accounts belong.

<table>
<thead>
<tr>
<th>Category</th>
<th>Accounts</th>
<th>Min</th>
<th>Q₁</th>
<th>M</th>
<th>Q₃</th>
<th>Max</th>
<th>Mean number of followers (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HCoastGuard, Diavgeia, GGNeasGenias, Govgr, Hellenicpolice, MinDevGR, GreeceMFA, Yyka_gov, YPEKA, OpenGovGr, Pyrosvesiti, PrimeministerGR, Synigoros, VisitGreecegr</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>13,499 (15,671)</td>
</tr>
<tr>
<td>2</td>
<td>Yp_Paideias, StartupGreece</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>13</td>
<td>3,910 (3,005)</td>
</tr>
<tr>
<td>3</td>
<td>DigiPlanGR</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>12</td>
<td>738</td>
</tr>
<tr>
<td>4</td>
<td>Gsis_gov_gr, PressParliament</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>621 (159)</td>
</tr>
<tr>
<td>5</td>
<td>Digitalgreece</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>469</td>
</tr>
<tr>
<td>6</td>
<td>GGMEofficial</td>
<td>2</td>
<td>2.7</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>76</td>
</tr>
<tr>
<td>7</td>
<td>e_ekpaideusi</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>77</td>
</tr>
<tr>
<td>8</td>
<td>Synigkatanaloti</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>38</td>
</tr>
</tbody>
</table>

In Table 2, it can be seen that in Category 1, accounts may be members of combinations of up to 13 accounts. That is, they may have common followers with up to 12 other accounts. Accounts in Categories 1 to 3 and, to a smaller degree, accounts in Categories 4 and 5 share followers with many other accounts. It is interesting that the majority of the accounts and all the accounts of the Ministries share followers with many accounts, since they belong in Categories 1 to 3. The accounts, GGMEofficial, E_ekpaideusi and Synigkatanaloti, are isolated.

In all the above cases, the number of followers of each account affects the ability of the account to share followers, meaning that accounts with more followers may share followers with many other accounts. Categories in Table 2 do not group accounts by means of similarity (shared followers), but because they have common distributions. It is a brief way to present the ability of accounts to share followers with other accounts. Belonging to the same category, two accounts do not necessarily share followers; they may share followers with other accounts. What is described here is that accounts of the same category share followers with equal numbers of accounts.

When using this procedure, accounts with many followers will be overrepresented and accounts with only a few followers will be underrepresented. For this reason and because of the fact that association rules result in distributions of followers and not groups of accounts sharing followers, in the next section, we develop a statistical approach to represent clusters of accounts sharing followers in a brief and compact way.

4.3. A Multivariate Statistical Approach

To develop a procedure for briefly grouping together accounts that share followers, we introduce the use of a similarity measure suitable for this kind of representation. The problem with choosing a similarity measure for Twitter accounts is that there is a great diversity of the number of followers of
the accounts. Accounts are not comparable for this reason. A correlation coefficient is not an adequate measure, because it considers both the unities and the zeros. The idea that we use in the present paper is that two accounts should be comparable despite their relative sizes, and only the unities should be counted. Therefore, when comparing two accounts, we considered the size of the smaller account and measure how many of its followers also follow the larger account. In Table 3, we present an example of how to elaborate on a suitable similarity coefficient [60].

Table 3. An example of the common distribution of followers for two Twitter accounts.

<table>
<thead>
<tr>
<th>Twitter account A</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td># common followers of A and B = a</td>
<td># followers of B, but not A = b</td>
</tr>
<tr>
<td># followers of A, but not B = c</td>
<td># not followers of A or B = d</td>
</tr>
<tr>
<td># followers of A = a + c</td>
<td># not followers of A = b + d</td>
</tr>
</tbody>
</table>

Suppose there are two Twitter accounts, A and B. Then, suppose that these two accounts have a followers in common. Table 3 shows that besides those a followers, there are also b followers that follow B, but not A and c followers that follow A, but not B, and of course, there are those followers (d), who follow none of the two.

The Simpson similarity coefficient S is defined as [60]:

\[
S = \frac{a}{\min(a + b, a + c)}
\]

S equals the number of common followers divided by the number of followers of the account that has the fewest followers. For example, Simpson similarity for the Hellenic police and Coast Guard accounts equals 0.774, because 77.4% of the followers of the Coast Guard account (1178 of 1522 followers) also follow the Hellenic police account (which has 31,423 followers in total).

A 27 by 27 symmetric Simpson similarities matrix is then constructed. This matrix serves as the input data in a multidimensional scaling procedure (MDS) [61]. This multivariate statistical procedure transforms original similarities into new ones in a space of fewer dimensions, in order to reduce the dimensionality. A goodness-of-fit index, named “stress” is used in order to decide whether the method reproduces the original similarities adequately. Stress is a loss function in this context, which is minimized using a procedure called “stress majorization”. Stress is a measure of the goodness-of-fit between the distances of the original data and the distances of the reproduced data. It measures the squared differences between ideal (low-dimensional) distances and actual distances in the high-dimensional space [59]. Small values of stress (lower than 0.05) indicate a good fit and representation of the original data to the new lower dimensional space. To achieve this value of stress, a reduction to four dimensions is made using the Simpson similarity matrix (Stress = 0.034). Therefore, for every account, four coordinates could now be used.

Next, a hierarchical cluster analysis (using the average linkage method) of the accounts when they are represented in this four-dimensional space produced six clusters. The decision on the exact number of clusters was made using a scree-plot of Wilk’s lambdas versus the number of clusters. Figure 6 presents a dendrogram of the hierarchical clustering and is used as a complementary source supporting the discussion of the findings. The dendrogram describes in detail how the accounts, forming a cluster, are actually joined to form the cluster. It shows what accounts are closer (according to the similarity
measure) than others and which accounts are grouped together before others join their group at a later stage.

Figure 6. Dendrogram for the hierarchical cluster analysis of the 27 accounts.

Cluster 1 is as follows: HCoastGuard, PressParliament, Hellenicpolice, GreeceMFA, Pyrosvestiki, GGMEofficial (mainly G2C, information provision); this cluster consists of the accounts of services and administrations of public order, the parliament and the press. e-Government services originating from the three main centers of state power are included in this cluster. From the dendrogram in Figure 6, it is obvious that Hellenic Police, Fire Brigade and, later, Coast Guard are joined to form a subcluster at an early stage. GGMEofficial and GreeceMFA are also joined together. These two primary sub-clusters are joined together and with PressParliament, at the last stage, to form Cluster 1. Therefore, Cluster 1 not only groups together the accounts of the administrations of public service, the parliament and the press, but it groups them at stages. At each stage, accounts of similar orientation or purpose are joined to a common subcluster.

Cluster 2 is as follows: Diavgeia, DigitalGreece, DigiPlanGR, Madecgr, Gsis_gov_gr, Govgr, HellasYAP, MinDevGR, YPEKA, OpenGovGr, StartupGreece, Yyka_gov, PrimeministerGR (mainly G2C, G2B, information provision, consultation, and collaborative environments).

The Prime Minister’s account, the Greek government account, the account of Ministries of Health and Environment and, especially, accounts concerning the promotion of open government and digital services are included in Cluster 2. It is interesting that governmental services regarding the promotion of digital and ICT services are included. The Prime Minister’s account is grouped with the Government account very early to form a subcluster. Furthermore, the HellasYAP and.opengovgr accounts are
grouped together. These two sub-clusters are joined with the Ministry of Development (MinDevGr). Diavgeia and Digitalgreece, while DigiPlanGr and StartupGreece are grouped together at an early stage. Next, these sub-clusters are joined with Yuka_gov, Gsis_gov_gr and Madcgr. In the same fashion as in Cluster 1, accounts in Cluster 2 are grouped together to form the final Cluster 2 in stages where accounts of the same or similar content, orientation or intended purpose are grouped together.

Cluster 3 is as follows: e_ekpaideusi, GGNeasGenias, Yp_Paideias, VisitGreecegr (mainly G2C, G2B, information provision). This cluster mainly includes accounts of services related to education and youth. E-ekpaideusi and Yp_Paideias, services concerning education, are grouped together. Furthermore, the cluster includes VisitGreecegr, which is about promoting tourism in Greece, and at an early stage, it is grouped together with GGNeasGenias, the Secretariat for youth.

Clusters 4 (Gsaeedu, G2C, G2B, information provision), Cluster 5 (Minfingr, G2C, information provision) and Cluster 6 are less populous, and they are mainly containing isolated accounts with only a few followers.

Cluster 6 (G2C, information provision, mediation) contains the accounts of two services. These services offer mediation between the citizens and the state (Synigoros) and between the consumers and the other stakeholders, such as suppliers and services (Synigkatanaloti).

It can be concluded that, to a certain extent, citizens follow groups of accounts in relation to their thematic similarity or orientation, for example, accounts regarding education or accounts regarding digital services, public order, etc., share followers. This is not a general pattern, but it seems to hold at large.

Clustering serves as a way to group together accounts that share followers as defined by the similarity measure. Clustering may group together accounts that belong to different distribution categories as those were presented in Table 2 and Figure 5. Table 2 presents groups of accounts, which have nearly the same distribution, not necessarily accounts that share followers.

5. Conclusions

The paper presented some quantitative features of the e-government Twitter accounts in Greece, regarding their followers and the fashion that they follow Twitter accounts. While it is interesting to record and measure the number of followers and the accounts they follow, it is also interesting to explore what are the combinations of accounts they choose to follow. By locating them, we can understand whether there are certain patterns of accounts followed, according to the citizens’ interests.

Two techniques are proposed to represent and describe such combinations of accounts that are followed by the same followers. The use of association rules relies solely on counting cases and provides an exhaustive representation and recording of several patterns regarding accounts with shared followers. To produce a brief representation of the accounts in clusters of those which have common followers, we introduce the use of a suitable similarity measure. Based on this measure, a cluster analysis is used to produce the clusters. MDS is used in a preliminary stage and is useful in the case that the data matrix is very big and dimensionality needs to be reduced.

Findings provide a snapshot of what accounts citizens choose to follow in order to stay informed on e-government matters through Twitter. Clusters and groupings of accounts reveal that, at large, citizens choose accounts of services that have common origins, orientations or content.
The paper answered the main queries posed in the beginning of Section 3:

1. It provided numbers of followers for each account. Some accounts are very popular regarding the number of followers, while others have only a few followers. Accounts of services regarding public order, the Prime Minister and promotion of Greek tourism are followed by many citizens;

2. While some citizens follow only one account, there are many citizens who follow several accounts. There exist groups of up to thirteen out of twenty seven accounts that share followers. To answer how many followers follow more than one account, the paper presented detailed counts of followers that follow just one or more than one account;

3. Accounts that share followers are presented in the findings of the association rules approach. The distributions of the number of accounts that every account shares followers with is displayed and discussed;

4. Cluster analysis produced groups of accounts that are followed by the same followers;

5. The analysis of the intended purpose and the content of the accounts that are clustered together show that followers choose which accounts to follow by means of the accounts’ thematic similarity. Accounts regarding education, public order and digital services are clustered together, since they share common followers. We may conclude that, to a certain extent, people choose to follow accounts of the same orientation, content or use.

Several previous studies, presented in Section 2.2 of this paper, have explored Twitter use in e-government, either at a local or a national level. They mostly use content analysis, social networking analysis and several Twitter metrics to explore what is the current status of Twitter use and how local and national authorities and individuals address citizens. In this paper, we considered the whole set of Greek government Twitter accounts addressing the citizens nationwide. We used a different approach to examine how citizens choose which account to follow, by examining how accounts share followers. A data mining technique was used to present the distribution of accounts that share followers, and a multivariate statistical technique resulted in locating clusters of accounts that share followers. To a certain extent, these accounts are similar regarding their scope, intended purpose and subject or target audience. The analysis followed an approach from the outside in, meaning that we first considered objective features of the accounts, such as the number of followers and, then, tried to examine what the similarities of the formed clusters are. The methodology adds to the literature on Twitter use in e-government by offering another, mainly statistical, complementary approach.

Future research will deal with the relation between followers’ activity in Twitter and the fashion by which they follow e-government Twitter accounts, along with the e-government account activity, as described by the average number of tweets and potential reach to citizens. Future research will also include the study of how the government uses the tweets in terms of content and information dissemination.

Author Contributions

The authors contributed equally in all the sections of the paper.
Conflicts of Interest

The authors declare no conflict of interest.

References


54. Crump, J. What are the police doing on Twitter? Social media, the police and the public. Policy Internet 2011, 3, 1–27.


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).