

Applying Apriori Rules Mining in Evaluating Digital Government Services Patronization by a Younger Generation of Users in Russia

Ebenezer Agbozo

Ural Federal University, Yekaterinburg, Russian Federation

Abstract – The Russian Federation has been on a gradual course of improving digital public service delivery using its Gosuslugi platform. Research into the patronization of services has not been explored. In order to assess services patronized by a younger generation of Russians, survey data from students in the Sverdlovsk region is analysed using standard statistical techniques and apriori algorithms and association rules mining technique to generate association rules. The results of the study are discussed as well as theoretical contributions of the study in the light of the research findings.

Keywords – Association Rules Mining, Russia, e-government, data mining, e-services usage, Gosuslugi, youth.

1. E-Services usage in Russia

The implementation of information communication technologies (ICTs) in the public sector giving birth to the concepts Electronic Government (e-government) and Digital Government has been of relevance to numerous nations [1]. Russia is one of such nations.

Kabanov & Sungurov [2] argue that technological levels hinder or influence e-government development and in Russia's case, for regions with low internet

penetration, the government and citizens do not have incentives to develop e-government beyond the minimal federal requirements. Kabanov & Sungurov [2] attribute the success of e-government diffusion in Russia to the crucial role played by the central government and its vertical influence (i.e. political, legal, organisational and financial) that compels regional authorities to comply with federal requirements, notwithstanding a lack of motivation to innovate among regional actors. Thus, the federal government of Russia seized the opportunity to ensure a positive trend of adoption of e-services by citizens and public service providers.

As at 2015, 13% of Russia's population were registered users of the official e-services portal (www.gosuslugi.ru) and e-services usage was reportedly low [3]. As visible in Figure 1., Russia's impressive path towards e-government development is expressed in the E-Government Development Index (EGDI) and Electronic Participation Index (EPI) from the United Nations Department of Economic and Social Affairs e-government survey. This outstanding performance is proof of a foundation built by presidential and public administrative reforms such as service standardization, monitoring government efficiency, multifunctional centers, and government portals for services, just to mention a few [4] that have played a role in creating the necessary environment for e-government adoption, progress and growth in Russia. In spite of the rise in development, Smorgunov [5] reveals that there has been high expenditure on ICT infrastructure, yet a slow development of e-government at the municipal level. Treakhtenberg and Dyakova [6] indicated that electronic services (e-services) integration within the public sector has increased and can be attributed to two factors: the presidential decree No. 601 (dated May 2012) which resulted in restructuring of service definitions and massive application of marketing approaches to increase service adoption and e-participation.

DOI: 10.18421/TEM84-14

<https://dx.doi.org/10.18421/TEM84-14>

Corresponding author: Ebenezer Agbozo,
Ural Federal University, Yekaterinburg, Russian Federation
Email: eagbozo@urfu.ru

Received: 10 July 2019.

Revised: 28 October 2019.

Accepted: 01 November 2019.

Published: 30 November 2019.

 © 2019 Ebenezer Agbozo; published by UIKTEN. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

The article is published with Open Access at www.temjournal.com

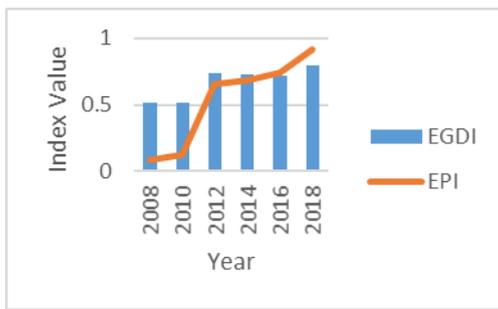


Figure 1. A decade of E-Government and E-Participation trend in Russia (data source: UNDESA¹)

Gosuslugi – which stands for state services, is the Russian state services portal which is “part of the infrastructure which provides data exchange and technological interaction of information systems used for rendering state and municipal services in electronic form”. The Public Services Portal of the Russian Federation or the Common Government Services Portal of Russian Federation as at 2015 had 22.5 million registered citizens [7]. In 2017, the Ministry of Communications and Mass Media indicated that 40 million Russians were registered users [8]. This indicates that a high percentage of Russians are embracing the public sector innovation in spite of the uneven nature of Russia’s e-government development [9]. Gosuslugi e-service categories are: Nature and ecology; Information; communication and advertising; Production and trade; Business, entrepreneurship, NGOs; Culture, leisure, sport; Work and employment; Security and law and order; Apartment, construction and land; Licenses, certificates, accreditations; Pension and benefits; My Health; Taxes and finance; Education; Transportation and driving; Passports, registrations, visa; and Family and Children².

Empirical studies have delved into factors that determine the usage of e-government services and Mensah et al. [10] concluded that demographic factors such as age play a role in willingness to use e-services. Styryn & Dmitrieva [3] encourage the development of capabilities to measure citizens’ usage of information services. Drawing from the literature above, this study aims at exploring e-service usage among young Russians (particularly university students) by answering the following research question:

¹ UN E-Government Knowledgebase | Country Data, URL: <https://publicadministration.un.org/egovkb/en-us/Data-Center>

² Gosuslugi Service Categories, URL: <https://www.gosuslugi.ru/category>

RQ1: What categories and combinations of public e-services of the "Gosuslugi" platform do young Russians use?

The next section discusses the data collection tools and techniques used in answering the research questions.

2. Data and Methodology

To answer the research question, the study gathered primary data using surveys. Survey questions were prepared using Google Forms in Russian language due to the linguistic capabilities of the desired target group and were assessed by an expert to ensure it in no way compromised on the privacy of the respondents as well as would be succinct. Using a purposive sampling technique, the survey questions were distributed among students in the Department of Systems Analysis and Decision Making of Ural Federal University in Yekaterinburg (Sverdlovsk region) of Russia. The Sverdlovsk region was selected since it is one of Russia’s most populated and cosmopolitan regions in Russia. As of 2013, with respect to data gathered by the Data of Ministry of Telecom and Mass Communications, 7.3% of the population of Sverdlovsk region were registered users of Gosuslugi [5]. In addition to that, Smorgunov [5] highlights Sverdlovsk as having a population of 4,316,852 with 7.3% registered users of Gosuslugi as of 2016. As a result of the youth being the main users of technology today [11], the study is based on the youth as its target group.

Between April and May 2019, the URL to the survey was distributed by social media and group messaging applications and furthermore distributed amongst other students who were colleagues to the initial recipients of the survey.

Of the ten (10) questions asked, aside demographic details of the respondents, only one is of relevance to this research:

- a. What categories of services of the "Gosuslugi" platform do you use?

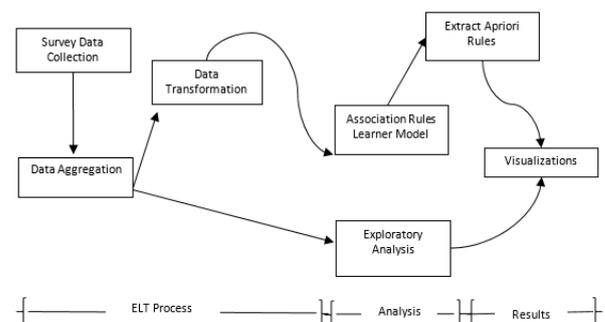


Figure 2. Methodological Research Model and its phases

Figure 2. illustrates the research path traversed to attain the goals of this study. Upon collection of the survey responses, the data was aggregated as a CSV (comma separated value) file and records were loaded into the RStudio IDE. The data was transformed from individual records to transactional data for the proposed analysis. Exploratory analysis was undertaken to identify schemes that understand the data, selection of a reasonable scheme which is followed by building a model to fit the data [12].

The study adopted the association rules data mining technique by building an apriori algorithm. Association Rules Mining (ARM) is essential in detecting unknown relationships which may also serve as the basis for decision making and predictions in areas such as telecommunication, banking, health care and manufacturing [13],[16]. Aside ARM's ability to reduce computation, I/O cost and memory requirement, Karthikeyan & Ravikumar [14] discuss various ARM approaches and their benefits. Drawing from Al-Maolegi & Arkok [13] and Hornik et al. [15], ARM is mathematically represented as follows with respect to this study's data (i.e. e-services):

$$I = \{i_1, i_2, \dots, i_k\}$$

Where **I** represents an item set of a records table with each row as transactions and items (i.e. e-services).

$$T = \{t_1, t_2, \dots, t_n\}$$

Where **T** is an expression of a transaction.

Thus, the association rule is defined as an implication in the form

$$X \Rightarrow Y, \text{ where } X \subset I, Y \subset I \text{ and } X \cap Y = \emptyset$$

The support which determines the frequency of the appearance of item sets (a set of e-services), is calculated as:

$$\text{Support}(X \Rightarrow Y) = \frac{|X \cup Y|}{n}$$

The lift which determines the percentage in which e-service Y is used with X, is calculated as:

$$\text{Confidence}(X \Rightarrow Y) = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)}$$

Thus, the study aims at discovering association rules, calculating the support and confidence with respect to the usage of e-services among Russian youth, primarily the Sverdlovsk region. Hence, ARM is the most suitable analytical technique for this study. The next section delves into the analysis results and discussions.

3. Results and Discussions

The research findings are summarized in Tables 1. and 2. as well as Figures 1., 2., 3., 4. and 5.

Table 1. evidently reveals demographic findings from the survey. Approximately 61% of the respondents were female; 74% were bachelor's students; 98% were of Russian heritage with the remaining hailing from former Soviet Union states (i.e. Ukraine, Georgia and Kazakhstan); and 94% resided in the capital of the Sverdlovsk region (i.e. Yekaterinburg) with the remaining living in surrounding cities such as Verkhnyaya Pyshma, Sysert, and Beryozovsky. Gosuslugi and the necessary e-services are available to foreign resident in Russia. The availability of Gosuslugi to foreign resident in Russia presents a great opportunity to ease the burden on foreign residents that require certain services such as police reports, vehicle registration and traffic fine payment. Majority of the respondents were between the ages 20 to 24, accounting for approximately 69% of the total sample size.

Table 1. Demography of survey respondents.

Demographic	Category	Value	%
Gender	Female	130	61.32
	Male	82	38.68
Education	Bachelor	157	74.06
	Specialist	51	24.06
	Postgraduate	3	1.42
	Master	1	0.47
Age	17 – 19	58	27.36
	20 – 24	147	69.34
	25 - 29	7	3.30
Citizenship	Russia	207	97.64
	Ukraine	3	1.42
	Kazakhstan	1	0.47
	Georgia	1	0.47
Residence	Yekaterinburg	199	93.87
	Other Cities	13	6.13

Figure 3. displays the absolute numeric frequency of e-services that young Russians from the survey sample size utilized in descending order. Evidently, the Passport, Registrations and Visa service category is the most resorted to. Hence, Figure 1. answers research question one (1) – i.e. ‘What categories of services of the "Gosuslugi" platform do you use?’.

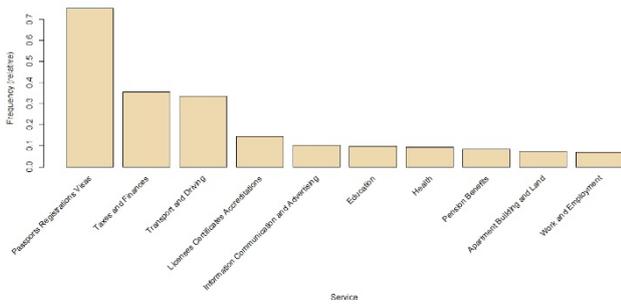


Figure 3. Relative frequency

The apriori algorithm calculated the support levels for four (4) instances – i.e. 10%, 5%, 1%, and 0.5% – in order to ascertain an appreciable number of association rules. From Figure 4., one can infer the following:

For a support level of 0.5%, the study identified an enormous number of rules for analysis purposes;

For a support level of 1%, an appreciable number of rules were obtained, also with a confidence level of approximately 50% - Hence, the chosen support level;

For support levels of 5% and 10%, there were a few rules available for analysis.

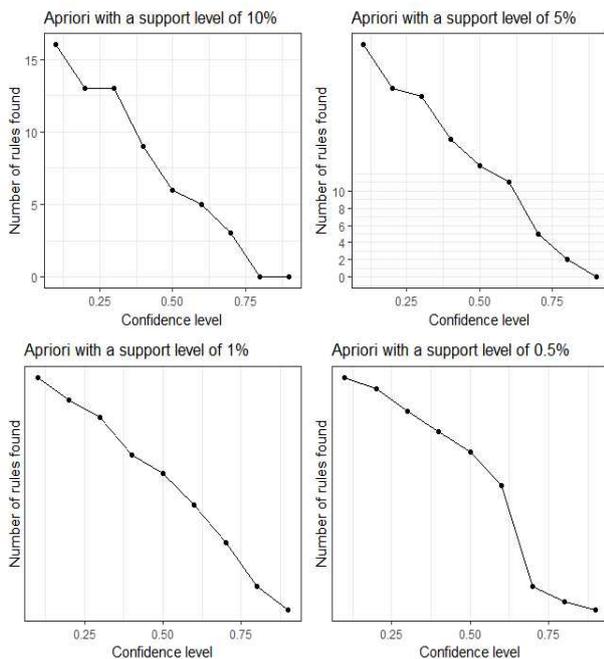


Figure 4. Apriori Algorithm with Varying Support Levels (10%, 5%, 1%, 0.5%)

Table 2. Association Rules Mining results

	Rules (LHS=>RHS)	Support	Confidence	Count
1.	{ } => {Passports Registrations Visas}	0.75233645	0.7523364	161
2.	{Pension Benefits} => {Passports Registrations Visas}	0.05607477	0.6666667	12
3.	{Education} => {Passports Registrations Visas}	0.08411215	0.8571429	18
4.	{Information Communication and Advertising} => {Passports Registrations Visas}	0.07009346	0.6818182	15
5.	{Licenses Certificates Accreditations} => {Taxes and Finances}	0.08878505	0.6129032	19
6.	{Licenses Certificates Accreditations} => {Passports Registrations Visas}	0.11214953	0.7741935	24
7.	{Transport and Driving} => {Passports Registrations Visas}	0.23831776	0.7083333	51
8.	{Taxes and Finances} => {Passports Registrations Visas}	0.24766355	0.6973684	53
9.	{Licenses Certificates Accreditations, Taxes and Finances} => {Passports Registrations Visas}	0.07476636	0.8421053	16
10.	{Licenses Certificates Accreditations, Passports Registrations Visas} => {Taxes and Finances}	0.07476636	0.6666667	16
11.	{Taxes and Finances, Transport and Driving} => {Passports Registrations Visas}	0.10747664	0.6388889	23

Table 3. Top 5 Association Rules and their interpretations

	Rule (LHS=>RHS)	Interpretation
1	Education} => {Passports Registrations Visas}	Approximately 86% of respondents who use the Education service also used the Passport, Registration and Visas service category
2	{Licenses Certificates Accreditations, Taxes and Finances} => {Passports Registrations Visas}	Approximately 84% of respondents who use both the Licenses, Certificates, and Accreditations service category and the Taxes & Finances service category also used the Passport, Registration and Visas service
3	{Licenses Certificates Accreditations} => {Passports Registrations Visas}.	Approximately 77% of respondents who use the Licenses, Certificates and Accreditations service also used the Passport, Registration and Visas service category
4	{ }=> {Passports Registrations Visas}	Approximately 75% of respondents used the Passports, Registrations and Visas service category only
5	{Transport and Driving} => {Passports Registrations Visas}	Approximately 71% respondents who use the Transport and Driving service category also used the Passport, Registration and Visas service category

Table 3. is interpretation of Table 2. and it reveals the top five (5) apriori rules and their interpretation of the results of analysis of the study from Figures 4. and 5. The results reveal that the most prevalent combination of e-services is the Education e-service category together with the Passport, Registrations and Visas e-service category.

The Passport, Registrations and Visas e-service category is made up of digitally available services to residents and citizens of the Russian Federation such as:

- Passport of the citizen of the Russian Federation,
- Foreign passport of the new sample,
- Registration and issuance of seafarer's identity cards to members of crews of ships, vessels of mixed (river-sea) navigation and fishing fleet vessels,
- Birth registration,
- Citizen registration,
- Death registration,
- Registration of adoption,
- Voluntary fingerprint registration,
- Providing addressing information,
- Invitations to enter the Russian Federation,
- Information and consulting services.
- The Education e-service category is made up of digitally available services to residents and citizens of the Russian Federation such as:
 - Recognition of foreign education certificates,
 - Verification of results of final certification,
 - Recognition of foreign medical education,
 - Admission to an educational institution,

- Birth of A Child (Receipt of documents for a child, Maternity leave, Paternity establishment, Receipt of a child's SNILS (i.e. pension document))
- Enrolment in an educational institution,
- Enrolment in kindergarten,
- Providing preliminary exam results,
- Recognition of foreign education documents,
- Recognition of foreign medical education,
- Enrolment in an educational institution,
- Information on current academic performance,
- District-specific educational services.

From the above list of services one can infer that these e-services are of value to students and citizens.

Graph-based visualizations are capable of offering an aggregated perspective on rules which are most important in any study and rules are represented as directed edges between item sets [17]. For the purpose of this study, Figure 5. illustrates the highest support levels.

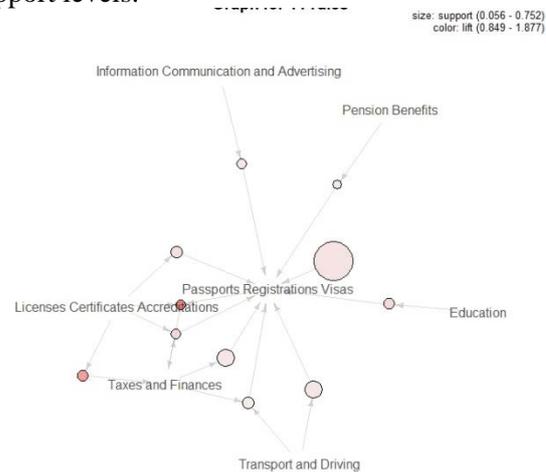


Figure 5. Graph for Top 11 rules

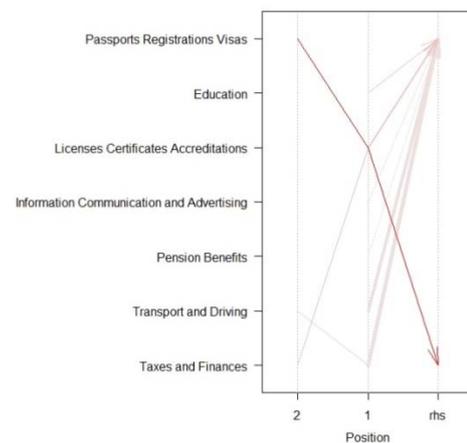


Figure 6. Parallel coordinates of Top 11 rules

Figure 6. illustrates a graphical representation of the parallel coordinates of the top 11 association rules – where RHS is the right-hand-side of the rule. In visualizing parallel coordinates, rules are represented as polylines joining the items in the antecedent followed by an arrow connecting another polyline for the items in the consequent [18]. This

graph visualizes which combination of e-services caused what kind of e-service usage. It is evident that Passports, Registrations and Visas e-service category is the most relevant in conjunction with other e-services among a majority of young Russian e-government users.

4. Conclusion

The study employed the data mining technique of apriori algorithms and association rules mining to extract association rules from survey data of young educated Russians – primarily university students in the Sverdlovsk region – and the e-services utilized. From the study, it is evident that students primarily use the Passport, Registrations and Visa service category and students who use the Passport, Registrations and Visa e-service also use the Education e-service categories.

One major limitation to this study has to do with the fact that data was collected from the Sverdlovsk region only. Future research activities will look into: widening the survey to obtain responses from other regions of Russia which will give a picture of young users and the e-services patronized per state.

With respect to this study's contribution to theory, insight is added to the body of knowledge within the sphere of application of apriori algorithms and association rules in decision making due to the fact that the results are relevant to policymakers in obtaining an overview of public sector services patronized by various age groups and as such build the e-services to support user needs.

References

- [1]. Peña-López, I. (2016). UN e-Government Survey 2016. E-Government in Support of Sustainable Development. Retrieved from: <https://ictlogy.net/bibliography/reports/projects.php?idp=3082> [accessed: 18 March 2019].
- [2]. Kabanov, Y., & Sungurov, A. (2016, June). E-Government development factors: evidence from the Russian regions. In *International Conference on Digital Transformation and Global Society* (pp. 85-95). Springer, Cham.
- [3]. Styrin, E., & Dmitrieva, N. (2015, November). Information services quality measurement: Russian federal authorities example. In *Proceedings of the 2015 2nd International Conference on Electronic Governance and Open Society: Challenges in Eurasia* (pp. 228-231). ACM.
- [4]. Литвинова, Т. Н. (2019). Развитие электронного правительства в России: проблемы и перспективы. *Право и управление. XXI век*, (3), 60-68.
- [5]. Smorgunov, L. (2016, June). From public administration reform to e-government: Russian path to digital public services. In *International Conference on Digital Transformation and Global Society* (pp. 232-246). Springer, Cham.
- [6]. Treakhtenberg, A., & Dyakova, E. (2019, February). Citizens and E-services: On-line Interaction with the Authorities in Russia. In *International Conference Communicative Strategies of Information Society (CSIS 2018)*. Atlantis Press.
- [7]. Dobrolyubova, E., & Alexandrov, O. (2016, June). E-government in Russia: meeting growing demand in the era of budget constraints. In *International Conference on Digital Transformation and Global Society* (pp. 247-257). Springer, Cham.
- [8]. Petrova, E. (2017, June). Common Government Services Portal of Russian Federation: How do People Use It? What is the Attitude Towards It?. In *International Conference on Digital Transformation and Global Society* (pp. 209-218). Springer, Cham.
- [9]. Kalyugina, S., Novikova, I., & Doryna, A. (2019, May). Leadership and e-government: a comparative analysis of the Republic of Belarus and Russian Federation. In *3rd International Conference on Social, Economic, and Academic Leadership (ICSEAL 2019)*. Atlantis Press.
- [10]. Mensah, I. K., Vera, P., & Mi, J. (2018). Factors determining the use of e-government services: An empirical study on Russian students in China. *International Journal of E-Adoption (IJEAA)*, 10(2), 1-19.
- [11]. Chvanova, M. S., Shlenov, Y. V., Anurieva, M. S., Kiseleva, I. A., & Molchanov, A. A. (2018, September). Internet-Socialization of Russian Youth. In *2018 IEEE International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS)* (pp. 598-601). IEEE.
- [12]. Grolemond, G., & Wickham, H. (2014). A cognitive interpretation of data analysis. *International Statistical Review*, 82(2), 184-204.
- [13]. Al-Maolegi, M., & Arkok, B. (2014). An improved Apriori algorithm for association rules. *International Journal on Natural Language Computing (IJNLC)*, 3(1), 21-29.
- [14]. Karthikeyan, T., & Ravikumar, N. (2014). A survey on association rule mining. *International Journal of Advanced Research in Computer and Communication Engineering*, 3(1), 2278-1021.
- [15]. Hornik, K., Grün, B., & Hahsler, M. (2005). arules- A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software*, 14(15), 1-25.
- [16]. Gayathiri, P., & Poorna, B. (2017). Effective Gene Patterned Association Rule Hiding Algorithm for Privacy Preserving Data Mining on Transactional Database. *Cybernetics and Information Technologies*, 17(3), 92-108.
- [17]. Hahsler, M., & Karpjenko, R. (2017). Visualizing association rules in hierarchical groups. *Journal of Business Economics*, 87(3), 317-335.
- [18]. Bruzese, D., & Davino, C. (2008). Visual mining of association rules. In *Visual Data Mining* (pp. 103-122). Springer, Berlin, Heidelberg.