



Bayesian analysis of internet access through apps as an e-government development strategy in Mexico

Rafael Rentería Gaeta y, Guillermo Alfonso de la Torre-Gea
Universidad De la Sierra Sur, Mexico

ABSTRACT

Mobile applications are types of technologies and resources connected to the Internet that constitute the backbone of mobile government, where its applications constitute a very broad development scenario. However, only a small number of mobile applications have managed to excel in e-government. Due to the complexity of the phenomenon, it is difficult to perform an analysis that takes into account each activity carried out on the internet as a variable of a system, so the objective of this study was to analyze the use of the Internet through Bayesian networks, that can be used for the development of strategies in matters of government. We managed to model a Bayesian network using the K2 classification algorithm, using the open data of the Intercensal Survey 2015 conducted by the National Institute of Statistics and Geography.

Keywords: *AI, mobile government, Digital Divided, Models*

Introduction

In the second decade of the 21st century, among the most important segments of information and communication technologies (ICT) worldwide are mobile telephony, Cloud Computing, and mobile applications. This type of technologies and resources connected to the Internet, constitute the backbone of a new paradigm in the public sector called a mobile government that is displacing or at least transforming the Electronic Government (e-Government).

In Mexico, the field of mobile applications is very broad; At the end of 2012, the largest segment was focused on complementary tools for mobile devices

(35%); games and entertainment followed (30%); the third site is occupied by applications used by small and medium-sized enterprises (SMEs) to market their products (25%); and the remaining 10% corresponded to other types of applications, including those used in the sector. In 2012, 65% of applications were free and 35% paid. Of the latter, 40% corresponded to audio and video applications; 31% were from the business world; 15% tools; 10% on health and welfare and the rest corresponded to games and entertainment applications. It also points out that the most used applications were Android with 42%, iOS with 40% and BlackBerry with 7% (AMITI-INFOTEC, 2013).

Currently, the mobile applications market generates thousands of millions of pesos in Mexico; It is worth mentioning that only in 2012 the download of applications generated 968 million pesos. There are thousands of companies throughout the country devoted to the design and development of mobile applications, but only a small number have managed to excel in the national market. However, the market is very dynamic and volatile as well as the modalities under which they are marketed; although many of them can be downloaded from the Web without cost, in reality, they are oriented to branding practices (creation and positioning of brands), marketing of products and services and generation of user databases.

The tendency of governments has been the adoption of mobile applications to expand their information channels and provide services to citizens. Mobile applications in the government sector can be classified into three segments: informative, interactive and transactional (Ramírez-Ramírez *et al.*, 2017).

The first segment is the largest and includes a set of applications aimed at providing information on public services, government actions and achievements, historical sites and information guides of various kinds, among others: the second segment, corresponds to all those applications that have a purpose to establish contact with citizens and provide them with a virtual space for participation and interaction; Finally, there are mobile applications that allow for tax payments and government services online, using the intermediation of a banking institution.

The UNIT (The Social Intelligence Unit) mentions that 2016 in Mexico increased the number of downloads of mobile applications of government by 45% with respect to 2015; The study was based on 75 mobile applications (apps) of dependencies and entities of the Federal Public Administration and State productive enterprises. In addition, it is mentioned that the average downloads of each application were 24,821 times during the year.

Sandoval *et al.* (2016) indicate that the highest growth of mobile applications during 2015 and 2016 occurred in the state (239%) and municipal governments (168%); To do so, he took a sample of 2013 applications from the three areas of government referred to. However, the data does not indicate that state and municipal governments present a greater penetration and maturity of mobile government; they only show that these areas of government are increasing the use of ICT and particularly that of apps.

Of the 290 apps located in Mexico, more than half are concentrated in three categories: economic promotion, tourism and culture (18%); justice, security and civil protection (17%) and education and health (15%). In contrast, the use of maps and location of sites (4%) and management of public services (10%) are the areas of the public sector where there are fewer available apps as shown in Figure 1.

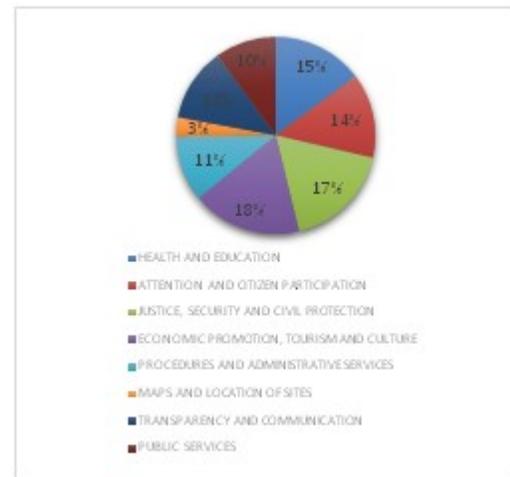


Figure 1: Categories of Government Apps in Mexico

Figure 1 shows a certain balance between the different categories, with the exception of the use of maps and location of sites. On the other hand, a study on the use and use of apps could offer a trend that could serve as a strategy for the development of mobile government. However, due to the complexity of the phenomenon, it is difficult to perform an analysis that takes into account each activity carried out on the internet as a variable of a complex system, so the objective of this study was to perform an analysis on the use of the Internet through Bayesian networks, which can be used to develop strategies in the field of electronic government.

Materials and Methodology

For the elaboration of this work, we made use of the open data of the Intercensal Survey 2015 conducted by the National Institute of Statistics and Geography (INEGI), from which those interested in the determination of Internet access were taken. The survey was applied to users, therefore, the variables corresponding to characteristics of access, use of social networks, purchase, and use of online banking, as well as interaction and use of electronic government, were selected. So, the variables to be considered were those that are available in Table 1.

Table 1: Variables analyzed in the study

Variables	Description
employ	He has done the internet job search activities
adults	You have visited adult sites
Encounters	He has carried out a search for online meetings
Mail	Send and receive mail
VoIP	Internet telephone conversations
Blogs	Create or visit blogs
Propel	Keep your own site
General	Get general information
Education	To support education
Banking	Perform online banking
Compras	Online shopping
Inst	Instant messages
Software	Software download
Gob_int	Interact with government
Gob-download	Download government formats
Gob_send	Fill or send government forms
R_socials	Access social networks
FB	Facebook
Twitter	Twitter
Inst	Instagram
Linked	LinkedIn
YouTube	YouTube
Whats	Whatsapp
Others	Other social networks
Music	Music
Health	Health
movies	Films

The analysis and study of each one of the variables considered was done through the Elvira software, which has its own format for coding models, a reader that interprets the coded models, a graphical interface that allows the construction of networks, what is intended for the edition and evaluation of probabilistic graphics models, specifically, Bayesian networks and flow diagrams (Pearl, 1988; Díez, 2004).

Bayesian networks (RBs) for Garrote *et al.* (2007) and Ropero *et al.* (2014) and are a statistical tool applied in environments under conditions of uncertainty. For this reason, Uusitalo 2007 (as cited in Ropero *et al.*, 2014) in this tool the most used solution is the discretization of the variables since this process involves the loss of statistical information by reducing it to the accuracy and precision in the resulting model. According to Ortiz-Vazquez *et al.* (2015), structural learning consists in obtaining relations of dependence and independence between the variables involved.

For this reason, one of the advantages of this type of network is that they simultaneously represent both the qualitative and quantitative dimensions of a problem (Hernández *et al.*, 2004, López, 2012). In addition, Bayesian networks identify interaction effects and models of nonlinear relationships between variables, likewise, it makes bi-directional inferences, that is, it goes from causes to effects and from these causes, it also makes adductive inferences to determine explanations of a phenomenon based on a set of evidence (Martínez *et al.*, 2018, Hernández-Ríos *et al.*, 2018).

Results and Discussion

We presented in Table 2 each of these variables to described, according to the relationships obtained once the Bayesian network was made, and the respective analysis of the incident tables, highlighting that the degree of relationship was first visualized of

the nodes of the network, by identifying the color, and direction of the generated arrows.

Table 2: Description of a priori and a posteriori probability of variables analyzed when the user interacts with electronic government in Mexico.

Variable	A priori (%)	A posteriori (%)
employ	0,61	0.9
adults	0.54	0.9
encounters	0.02	0.01
mail	0.78	0.99
voip	0.21	0.08
blogs	0.54	0.92
own	0.07	0.03
general	0.92	0.99
education	0.31	.09
bank	0.07	.05
purchases	0.07	.05
software	0.6	0.94
gob int	0.48	1.00
gob-download	0.51	0.96
gob_send	0.49	0.95
r_socials	0.84	0.98
fb	1.0	1.0
twter	0.1	0.05
inst	0.09	0.04
linked	0.01	0.01
YouTube	0.83	0.97
whets	0.87	0.98
others	0.01	0.0
music	0.63	0.93
health	0.69	0.97
movies	0.09	0.05

Subsequently, the analysis of the incidence tabs was performed according to De la Torre-Gea *et al.* (2014) and the relationships obtained in the first step, since in the second stage the a priori probability of each of the variables is identified, as well as the a posteriori probability that can be identified as they change or interrelate a given variable that another one has or has happened.

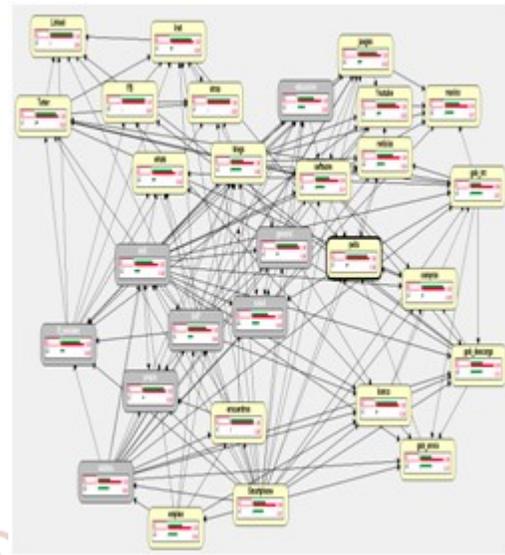


Figure 2: Bayesian network and inference of the use of the internet when the user interacts with e-government in Mexico

Figure 2 shows the relationship between each of the analyzed variables, highlighting that the variable with the highest ratio is the use of Smartphone for Internet access, which corresponds to 46% of the sample, that is, with The use of mobile government services depends on this variable and its respective percentage.

The user profile of the population that uses the internet in Mexico to carry out activities related to e-government mainly shows adults, who use social networks, making use of information seekers, mainly in topics such as education, employment, and health, in addition, to make use of email and VoIP. They are users with capabilities that place them as belonging to an information and knowledge society, where there has not been a digital divide and, therefore, they have the necessary equipment and knowledge to function adequately in that society.

Therefore, the way to increase the number of users in the e-government through mobile applications would be related to increasing the participation rates of the lagged population in the use of ICT, mainly by increasing the coverage of Internet access, as well as training adults who do not use these technologies.

Conclusion

The mobile applications market has increased in recent years, however, this growth is not reflected in the number of mobile government users. The population with capacities that place them as

belonging to an information and knowledge society, where there has not been a digital divide, corresponds to the user profile of the e-government. The e-government web portals are a source of a large amount of open data. The analysis of sayings through the use of artificial intelligence tools such as Bayesian Networks and Classifiers are very useful for decision making as an e-government development strategy in Mexico.

References

1. Alvarado, J. A. (2013). Modelos de monetización de aplicaciones móviles: análisis del caso meet2go. Universitat de Barcelona.
2. AMITI-INFOTEC. (2013b). Estudio de perspectivas y estrategias de desarrollo y difusión de aplicaciones móviles en México. Investigación Documental.
3. Botello, Peñaloza H. A. (julio- diciembre, 2015). Determinantes del acceso al Internet: Evidencia de los hogares del Ecuador. *Entramado* 11(2), 12-19.
4. De la Torre-Gea, G., Delfín-Santisteban, O., Torres-Pacheco, I., Soto-Zarazúa, G., Guevara-González, R., & Rico-García, E. (2014). Redes bayesianas aplicadas a un modelo CFD del entorno de un cultivo en invernadero. *Agrociencia*, 48(3), 307-319.
5. Díez F. J. (2004). Teaching probabilistic medical reasoning with the Elvira Software. *Research and Education*. España: Dept. Artificial Intelligence Universidad Nacional de Educación a Distancia.
6. Garrote, L., Molina, M., & Mediero, L. (2007). Probabilistic forecasts using Bayesian networks calibrated with deterministic rainfall-runoff models. *Extreme hydrological events: new concepts for security*, 173-183.
7. Hernández, J.; Ferri, C.; Ramírez, J. (2004). *Introducción a la minería de datos; Capítulo 10: "Métodos Bayesianos"; PEARSON EDUCACION*.
8. Hernández-Ríos, María Esther; Fernández-Tapia, Joselito y De la Torre-Gea, Guillermo. (2018). Bayesian analysis as a method to determine the limitations and advances of e-Justice. *International Journal of Trend in Scientific Research and Development*, 2(2); 117 - 122.
9. Instituto Nacional de Estadística y Geografía (2015). *Encuesta Intercensal, 2015*. México: INEGI.
10. López Puga J. (2012). *Cómo Construir y Validar Redes Bayesianas con Netica*. *Revista Electrónica de Metodología Aplicada*. 17 (1). 1-17.
11. Martínez, Elia y De la Torre-Gea, Guillermo. (2018). Analysis of the ICT user profile for e-government through Bayesian networks. *International Journal of Trend in Scientific Research and Development*, 2(2): 97 - 102.
12. Ortiz-Vazquez IC, Pérez-Robles JP, Fernández-Loyola R, Pérez-Brito JF, De La Torre-Gea GA. A multivariable computational fluid dynamics analysis method based on Bayesian networks applied in a bioreactor. *Journal of Applied Chemical Science International*. 2015;6(1):10–17.
13. Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. Morgan Kaufmann, San Mateo, CA.
14. Ramírez-Ramírez, R., De la Torre-Gea, G.A. (2017). Analysis of Public Information Requests in e-government through Bayesian Network models. *Journal of Scientific and Engineering Research* 4 (12), 424-432.
15. Roper, R. F., Aguilera, P.A. Fernández, A., Rumí (febrero, 2014). *Redes bayesianas: una herramienta probabilística en los modelos de distribución de especies*. *Ecosistemas* 23(1), 54-60.
16. Sandoval, R., Marín, J. & Romero, T. (2016). *Reporte del Ranking de Portales Estatales de Transparencia de México 2016*.