Chapter 3
A Fuzzy-Based Discounts Recommender System for Public Tax Payment

Jaime Meza, Luis Terán, and Martha Tomalá

3.1 Introduction

Taxes are the most common source of government incomes, which has influence in achieving socio-economic, political, and macroeconomic objectives of countries and other types of territorial divisions. Taxes are a legal instrument for increasing resources into the government to enhance its economic development. In the work of [25], the author highlights that tax payments are a major source of income of governments and are considered a fiscal instrument for regulating and resolving economic and social policies. At the same time, taxes are considered a mechanism for enhancing economic growth. In the case of this work, the tax payment historical behavior in Ecuador is analyzed. In Ecuador the tax payments are fundamental to financing the government national. Unfortunately, Ecuadorians' tax payment culture can be consider as low with high levels of evasion in different sectors of the economy [3, 31]. However, in spite of historical citizens' behaviors, nowadays; cultural conditions seems to be changing. According to statistics reports of the Internal Rents Service (from Spanish, Servicio de Rentas Internas[1] (SRI)), in the

[1]SRI: http://www.sri.gob.ec/web/guest/estadisticas-generales-de-recaudacion.

J. Meza (✉)
Universidad Técnica de Manabí (UTM), Portoviejo, Ecuador

L. Terán
University of Fribourg, Fribourg, Switzerland

Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador
e-mail: luis.teran@unifr.ch

M. Tomalá
Municipality of District of Quito (MDMQ), Quito, Ecuador

A. Meier et al. (eds.), *Applying Fuzzy Logic for the Digital Economy and Society*, Fuzzy Management Methods, https://doi.org/10.1007/978-3-030-03368-2_3

last years (2005–2016), the incomes ratio in Ecuador has been increased by 29%. This growth rate is due to public policies in taxes matter in Ecuador. Several of these policies are related to applying information and communications technologies (ICT) tools within taxes processes.

According to the European Commission[19], *eGovernment* is defined as a way to provide a wide variety of benefits including efficiency and savings for governments and businesses, transparency, and improvement of citizens' participation in political issues. One of the main challenges for the development of *eGovernment* is related to the implementation of ICT. However, *eGovernment* is much more than just ICT, it also involves rethinking organizations and processes and changing the behavior of public administrators so that public services can be delivered more efficiently to citizens.

eCommerce emerged due to the influence of the ICT revolution. Together with *eCommerce*, RSs were introduced as a mechanism to increase business incomes [35]. In the same way, several researchers have pointed out the benefits of applying the best practices of RSs to *eGovernment* solutions [26, 27, 38]. In spite of the efforts presented, the use of RSs as way to increase citizens' awareness towards improving tax payments is an application that needs more attention.

The Municipality of Quito (from Spanish, Distrito Metrolopitano de Quio (DMQ)) faces problems with a past due portfolio that includes a big group of citizens. According to report DAI-AI-0026-2017 presented by the Comptroller General of Ecuador, the past due portfolio until December 2015 was 31,284.41092 USD [10]. This trend continues generating delays in the execution of public services. On the other hand, the Municipality of Quito does not have processes and systems to setup discounts focused on citizens. Discount methods are mainly designed regarding payment deadlines. This has become in a big problem for the city. Moreover, every 2 years the most important taxes (urban and rural) are constantly changing. This process is developed and focuses on construction but it is not applied to citizens. This model request several legal reforms to change the discount formula. The constant modifications in the laws have an negative impact in the trust of citizens and it reduces their engagement level with the municipality [13, 14, 17].

These issues pushed tax authorities inside the Municipality of Quito to apply alternative and creatives approaches to reduce and mitigate the effect of a past due portfolio and motivate citizens towards improving their payment behavior [15, 16, 33]. Applying personalized discounts centered on citizen payment behavior with the use of RSs within the Municipality of Quito, could help to improve the issues mentioned above. The proposed methods in this chapter could present the following benefits:

- Improving income taxes.
- Reducing the due portfolio.
- Supporting decision-making processes to create/update taxes.
- Increasing citizens' engagement.
- Fostering the usefulness of RSs as a way to develop eGovernment.

Next sections are structured as follows. First, Sect. 3.2 gives a brief introduction and description of tax payments, RSs for *eCommerce* and *eGovernment*, and fuzzy logic, which are the basis of the model proposed. Then, Sect. 3.3 presents the methods used by the recommendation system model proposed. Additionally, in Sect. 3.4 a simulation model is presented. Finally, the authors give their concluding remarks and the future outlook in Sect. 3.5.

3.2 Background

3.2.1 Tax Payments in Ecuador

According to the Organization for Economic Cooperation and Development (OECD), the term *taxes* should be understood as compulsory, unrequited payments to general government, which means that the benefits provided by governments to taxpayers are not normally in proportion to their payments [30].

The current tax payment system in Ecuador is based on the following principles: legality, generality, equality, proportionality, and non-retroactivity [36]. The Ecuadorian tax system is based on indirect and direct taxes. In 2015, the fiscal pressure in Ecuador was around 21.7% [9]. On average, taxes represented 76% of the Ecuadorian GDP in 2016 [7] and are administrated by the SRI. On the other hand, external taxes are administrated by the National Customs Service of Ecuador (from Spanish, Servicio Nacional de Aduana del Ecuador[2] (SENAE)), which also has nationwide jurisdiction. Local and provincial taxes are created and administered by each local government as municipalities and provincial councils, respectively.

Municipalities in Ecuador use different discount methods as way to persuading citizens to pay their taxes. The most common method is to offer discounts by early payment [5].

3.2.2 RSs for eCommerce and eGovernment

RS techniques are applied in several domains. In the same way, Schafer et al. [35] show evidence that RSs can increase businesses' incomes. Additionally, the survey developed by Lu et al. [27] presents the state of the art in recommender system applications. A list of domains were evidenced with positive outcomes, which includes fields such as: *eGovernment*, *eCommerce*, *eLibrary*, *eLearning*, *eBusiness*, *eTourism*, *eResources*, and *eGroups*, among others.

In the academic literature, the most used mechanisms to develop RS applications are: collaborative filtering (CF), content-based (CB), and hybrid [8]. The main

[2]SENAE: https://www.aduana.gob.ec/.

problem in the CF approach is related to the so-called sparsity problem, when the number of items rated is small compared to the total number of items [38]. On the other hand, CB faces the problem of overspecialized recommendations [27]. To solve these problems, many advanced recommendation approaches have been proposed including social network-based RSs, fuzzy-based RSs, context awareness RSs, and group RSs [27].

In the work of Yager [41], the author points out that an important component of CF that is the calculation of similarity of interest based on correlations between individuals for predictions and recommendations. Several issues have been identified when using CF approaches, based on this issue. For that reason, a new approach so-called reclusive approach, was proposed. It is based upon finding a similarity between objects while the CF approach is based on similarity between people. This approach considers an RS as a collection of objects as, $O = \{o_1, \ldots, o_n\}$ and it recommends to a user objects of O that could interest him. In [41], the proposed method considers five issues: RS as object collections, object representation, user preferences modeling, user profiles, and environment.

The basis of this approach is an object and its membership degree. According to Zadeh [43], an object do not have a defined membership criteria. For instance, in the case of the class *animal*, it has a clear definition (e.g., dog, cat, horse, etc.). However, the class *people's ages* (e.g., baby, boy, young, and old), an imprecise decision emerges. In this arena, where the decision-making process is not obvious, fuzzy sets and fuzzy logic appears as a solution.

3.2.3 Fuzzy Logic Overview

In the work of Zadeh [43], the author pointed out that fuzzy sets provide a natural way of dealing with problems in which the source is imprecision. On the other hand, in sharp logic, the values obtained after executing a sentence are binary (true, false). These results are represented by *crisp sets*; however, several solutions in real life need values in the interval [0, 1]. In this case, fuzzy logic is similar to human decision-making with its ability to work with approximated data to find precise solutions [1].

Fuzzy set theory was introduced by Zadeh in 1965 [43]. It described the concept of membership degree. *Crisp sets* behaviors are denoted by: If x belongs to a set A then $x \in A$, otherwise $x \notin A$. Therefore, for each x of the set A there are only two responses: either x is an element of A or it is not. Using the concept of membership degree, each element can be represented by a function that defines the values allowed. The values allowed are drawn as linguistic values. A linguistic value refers to a label representing knowledge that has meaning determined by its degree of membership function. For instance, $x_1 = old$ with the degree $\mu = 0.8$ means that the variable x_1 has a linguistic value represented by the label *old*, whose meaning is determined by the degree 0.8.

To illustrate this concept in the case presented in this chapter, a *crisp set* for the linguistic variable $discount$ as a function of the payment date of a citizen is used. It tries to resolve the following problem: A local government institution wants to create a discount over one of its taxes. The discount strategy expects that citizens pay as soon as possible to maximize their annual budget. In represents that if one citizen pays between the discount period he gets a discount, but, if he pays in the first days he gets a gradual discount.

The class *discount* storage the discount periods in days. It is defined as follows: $discount = \{1, .., 180\}$. The characteristic function presented in Eq. (3.1) has only two values: true (ten) or false (zero).

$$X_{discount}(x) = \begin{cases} \text{if } 1 \leq x \leq 180 & 1 \\ \text{if } 180 < x < 366 & 0 \end{cases} \tag{3.1}$$

The function presented in Eq. (3.1) is not useful for the calculation required by the local government institution, since it does not allow setting a real discount value according to payment date.

This issue can be resolved using fuzzy sets with membership functions. Zadeh gives a number for each value inside the universe set; this number is the degree in which the element is inside the set. For instance, in the case of the set *discount*, there are some values that represent the citizen payment date; therefore, different degrees of *discount* could be represented.

Zadeh defines fuzzy sets as A in X is characterized by a membership function $f_A(x)$ that associates each element in X with a real number in the interval [0, 1], with the value of $f_A(x)$ at x representing the grade of membership of x in A. Equation (3.2) shows the membership function for the fuzzy class *discount*. The equation presented in Eq. (3.2) responds to a gamma inverse fuzzy function definition.

$$X_{discount}(x) = \begin{cases} \text{if } 180 < x < 366 & 0 \\ \text{if } x \in (1, 180) & 1 - ((x-1)/(180-1)) \\ \text{if } x = 0 & 10 \end{cases} \tag{3.2}$$

The level of granularity used in Eq. (3.2) allows the local government institution to apply a real discount related on payment date. The real values after and before could be represented in linguistic variables by a collection of quantifiers such as: near, close, approximately, etc. [44]. Finally, fuzzy set theory uses a large volume of operations and properties that can be considered as tools to apply in different scenarios.

3.2.4 Fuzzy Logic Applied to Marketing

In the work of Donzé and Meier [18], the authors define marketing as way to identify and pick up the customers' needs. They also point out the use of customer relationship management (CRM) as a strategy for building customer equity and improving financial revenue. In this sense, fuzzy logic can be applied to marketing in a wide set of tasks.

eGovernment is defined as a subset of the exchange relationships of *eCommerce*. Therefore, *eMarketing* is a way to improve the relations between citizens and government institutions through ICT tools. In the work of Meier and Stormer [28], authors propose to add *eMarketing* as an element into the value chain of *eBusiness*.

Several researchers highlighted the importance of fuzzy logic in both *eCommerce* and *eGovernment* as an effective means of decision support, not only from a business perspective but also from the clients or citizens perspective. Nowadays, fuzzy classification models are used in the implementation of RSs. In the work of Lu [27], the author highlighted a number of RS approaches and the techniques implemented. A summary is presented in Table 3.1.

Additionally, other authors have presented positive outcomes on the use of RSs in *eCommerce* and *eGovernment* [11, 12, 22, 23, 26, 27, 39, 42]; however, the application of RSs as a way to increase the citizens' awareness towards tax payments is poorly explored.

Table 3.1 RSs techniques applied, adapted from [27]

	Feature		
Name	*Domain*	*Technique*	*Period*
Smart participation [37]	*eGovernment*	Fuzzy clustering	2014
TPLUFIB-WEB	*eGovernment*	Fuzzy linguistic modeling, hybrid, CB, CF	2014
Methods for therapy [21]	*eHealth*	CF, Demographic-based RS	2016
Smart BizSeeker [27]	*eGovernment*	CF, hybrid, fuzzy sets	2013
Procurement [45]	*eGovernment*	Fuzzy logic, item-based CF, Bayesian approach	2015
Dissemination of information in university digital libraries[32]	*eLibrary*	Hybrid, 2-tuple fuzzy linguistic approach	2017
Proactive and reactive e-government services[6]	*eGovernment*	Hybrid with ontology-based recommendation model	2015
RS for elderly people [34]	*eHealth*	Hybrid	2015
Personalized E-learning [40]	*eLearning*	Hybrid, fuzzy tree and learner profile	2015

3.3 Fuzzy-Based RS Model

The model presented in this chapter is centered on citizens and their payments behaviour. Therefore, the reclusive approach for recommendation presented by Yager [41] is applied.

Based on the targeted marketing definition as results of RSs in *eCommerce* field, the approach proposed uses the definition of *discount* as a way to persuading citizen to paying taxes. In this way, it creates personalized discounts to each citizen. The model is set by two types of discounts centred on citizens, general and specific, defined as follows.

- General discount—It is applied in relation to the historical payment behaviour.
- Specific discount—It is related to specific issues that e-government institutions could be interested (e.g., dead line, risk relevance, etc.)

Both, general and specific discounts have been implemented using fuzzy sets definitions. Such definitions were mixed in a Euclidean space to set specific discount for specific citizens, where each discount percentage is unlikely. In the next section, the fuzzy sets used and its details are shown.

3.3.1 Fuzzy Sets

This section aims is to shows every fuzzy definition used to implement the discount methods but also the recommendation notification policies.

3.3.1.1 General Discount

To implement the general discount, the proposed model uses the historical citizens behaviors considering the number of times that the citizen was included in some payment condition as it is show in Table 3.2.

Payment conditions are applied regarding a group of taxes in the past. The values are computed using the fuzzy classification sets presented in Fig. 3.1. The

Table 3.2 Condition of payment by taxes

	Feature		
Condition	Rating (R)	Lexical meaning	Target taxes
Discount payment	+++++	Excellent	All
Non-discounted payment	++++	Good	All
Penalty payment	+++	Average	All
Payment after notification	++	Regular	All
Payment after judicial notice	+	Bad	All

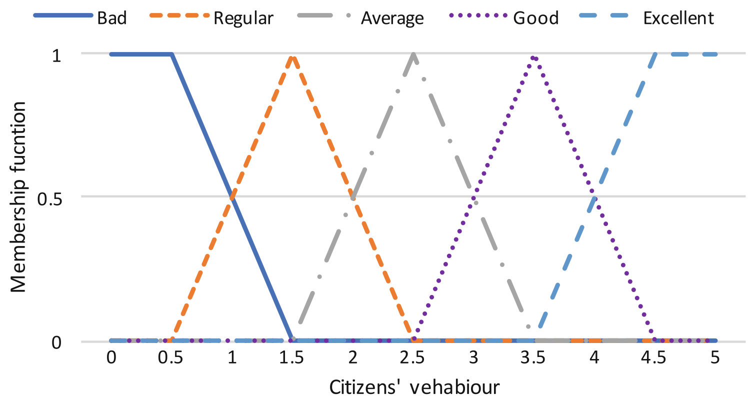

Fig. 3.1 Fuzzy sets for citizens behavior

Table 3.3 Discount by rating

Linguistic value	Discount %	Rating
Excellent	81–100	3.5–5.0
Good	61–80	2.5–4.5
Average	41–60	1.5–3.5
Regular	21–40	0.5–2.5
Bad	0–20	0–1.5

combination presented in Fig. 3.1 allows us to set a citizens classification, as well as a membership degree, defined as citizen behavior ranking (CBR).

Five fuzzy sets were drawn (Fig. 3.1) over the linguistic variable *citizens behavior*. The linguistic values are related to the citizens payment behavior (i.e., bad, regular, average, good, and excellent). Each linguistic values have a corresponding discount by crisp group as shown in Table 3.3. The ranking values shown in Table 3.3 have a corresponding fuzzy set (refer to Fig. 3.1); however, each ranking value is overlapped by other value. Therefore, the discount percentage does not belong to an element in one category only. In this scenario, the final discount will be done by membership degree on each group of fuzzy sets and the range of discounts set as shown in Table 3.3. This result is named citizen behavior discount percentage ($CBDP$) as shown in Eq. (3.3).

$$CBDP = \frac{\sum_{i=1}^{n} \{|(RF - (RI - 1))| \times \mu_i\}}{n} \tag{3.3}$$

where RF is the final value on the column discount in % as shown in Table 3.3. RI is the initial value on the column discount in % in Table 3.3. μ_i is the membership degree of the fuzzy set i. Finally, n numbers of fuzzy sets overlapped.

$CBDP$ computes the discount percentage by every group of taxes to compute the global percentage discount by citizen ($GCBDP$). Each group of taxes is assigned a weight (W) as shown in Eq. (3.4).

$$GCBDP = \sum_{I=1}^{N} W * CBDP \tag{3.4}$$

where, N is the total number of group of taxes by citizen. W is the ratio between the value to paying of tax group and the total debt for the citizen. As an example, citizen A has to pay three taxes: urban, patent, and operation. His total debt is 100 and the values to pay by each group of taxes are 80, 10, and 10. Therefore, the value of W assigned to each group of taxes are: $urban = 0.8$; $patent = 0.1$, and $operation = 0.1$. Finally, $CBDP$ is the individual discount percentage by group of taxes.

3.3.1.2 Specific Discount

The general discount presented in the previous section is the key element in the citizens payment behaviour. It could be mixed with specific discounts related with other factors (e.g., dead line, risk relevance, etc.). Specific discounts percentages (SDP) are computed as ratio between a point inside of Euclidean space and the maximum distance value (MDV) that can take negatives or positives values of the scale. It is shown in Eq. (3.5).

$$SDP = \frac{\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2)}}{MDV} \tag{3.5}$$

Both general and specific discounts are added for getting the final discount (FD) by group of taxes. It is shown in Eq. (3.6).

$$FD = \frac{(CBDP + SDP)}{2} \tag{3.6}$$

According to previous definitions, the model proposes two specific discounts. The first one is related to deadline payment, and the other with the risk relevance. Both discounts are combined with CBR and plotted on an Euclidean space as shown in Fig. 3.2. To compute the discounts according to Fig. 3.2, consider the X axis in region dead line represents time, and risk relevance is the ratio between citizens debt and the citizens with the most debt inside each group of taxes; moreover, points labeled as $X1 - Y1$ (as shown in Fig. 3.2) are the initial points to compute the Euclidean distance.

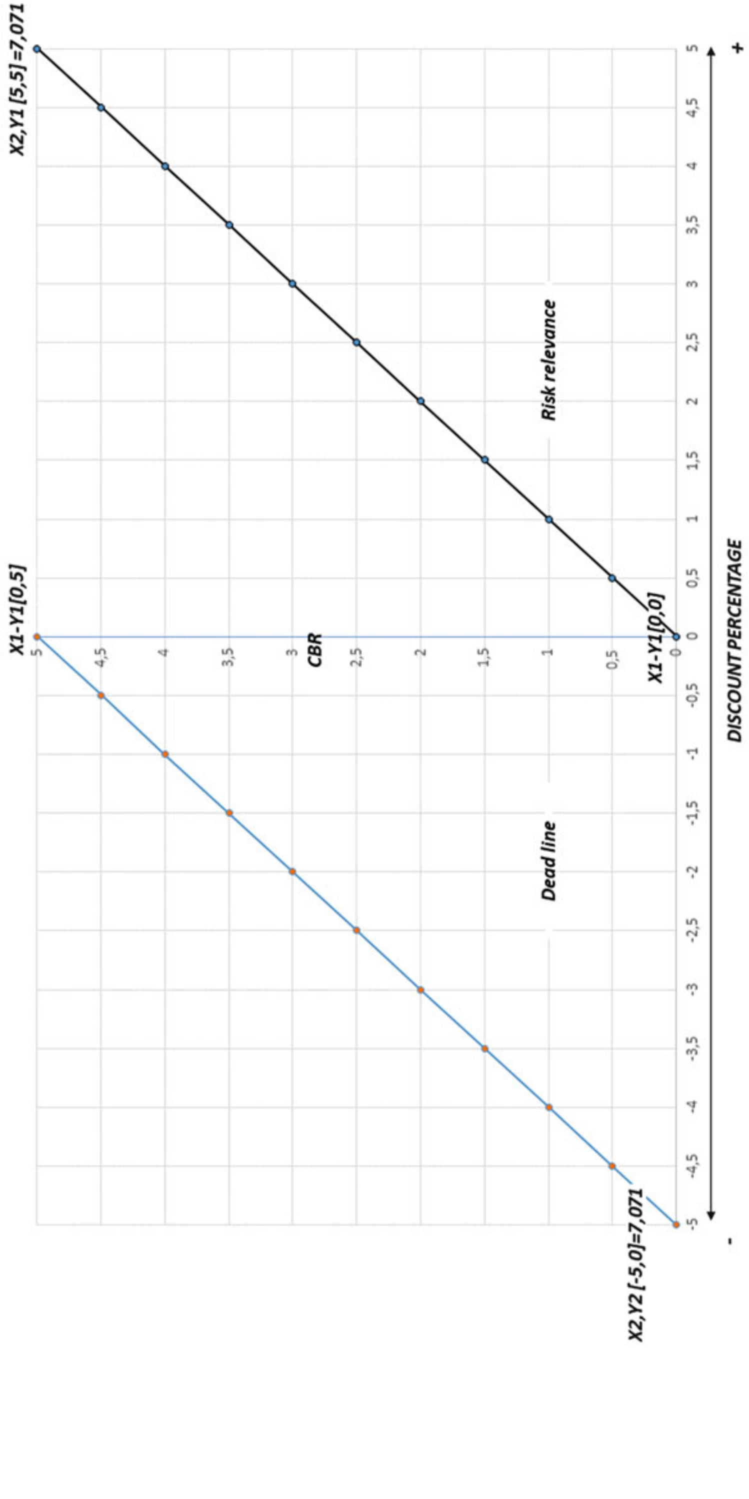

Fig. 3.2 Discounts combined with CBR

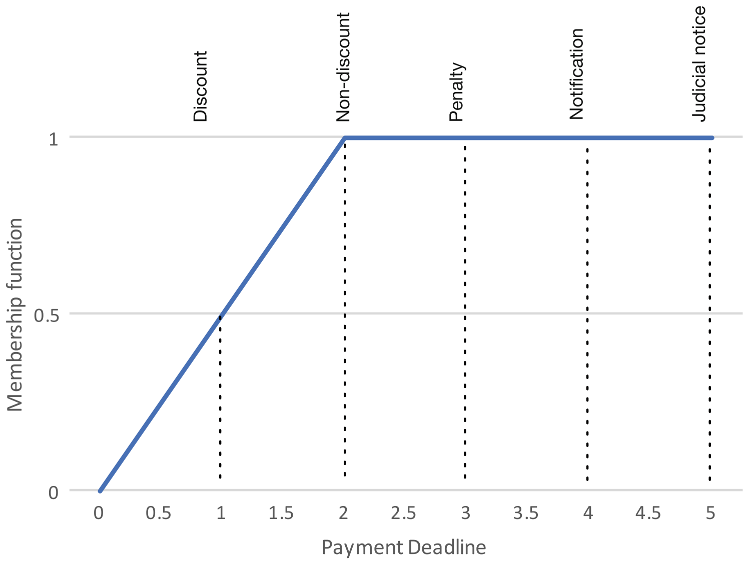

Fig. 3.3 Fuzzy sets by payment deadline

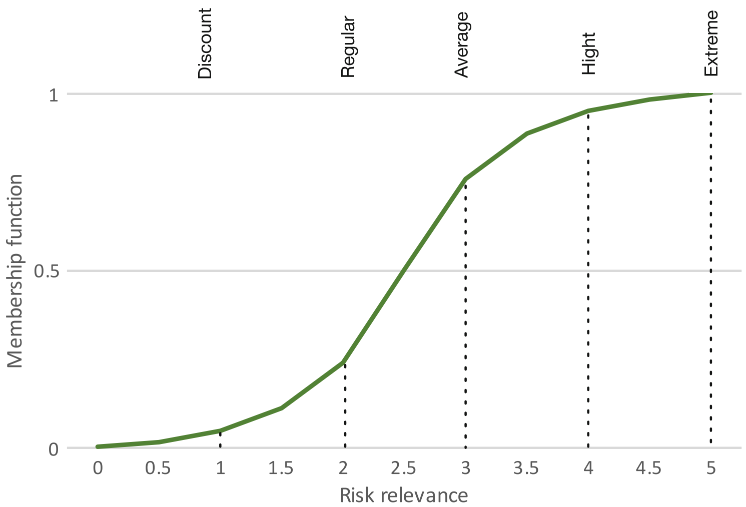

Fig. 3.4 Fuzzy sets by risk relevance

3.3.1.3 Recommendation Notification Policies

The fuzzy sets shown in Figs. 3.3 and 3.4 allow to setup the membership degree to define different recommendation policies, for example, if the membership function is 1 in both scenarios the recommendation should be sent to citizen every standard time.

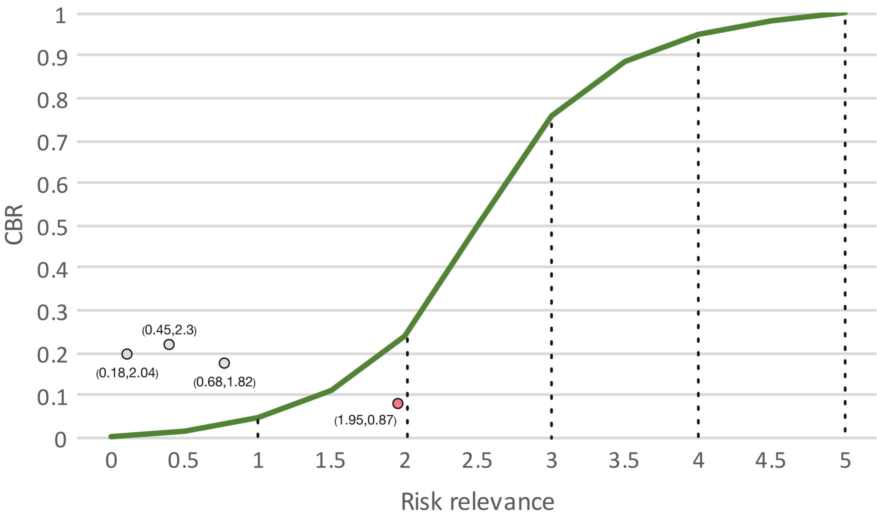

Fig. 3.5 Fuzzy profile (FP)

Citizens behavior fuzzy sets allow establishing recommendations centered on citizen. Therefore, the proposed model combines every fuzzy set (payment deadline and risk relevance) in an Euclidean space to compute the centroid defined as fuzzy profile (FP). Figure 3.5 presents the fuzzy sets of the citizens behavior and risk relevance combined in an Euclidean space, it is computed using Eq. (3.7). The FP is computed by setting up the initial centers as $(0, 0)$. The X axis represents the relevance (or deadline) and the Y axis is the point generated by CBR as shown in Fig. 3.5.

$$FP = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{3.7}$$

The FP presented in Fig. 3.5 generates a recommendation point with value 0.87. The risk relevance is calculated considering the value five for the citizen with the highest tax debt; therefore, the relevance for a specific citizen is the ratio between his or her debt and the citizen with the highest debt.

This information changes in *real time for each citizens payment*. On the other hand, the deadline for payment relevance is calculated considering the value two (non-discounted payment) as the user is moving towards the risk zone. Therefore, in this case notifications will be send. The notifications regarding citizens payments will depend on their behaviors. Both risk relevance and deadline payment generate a *recommendations matrix* to save recommendations messages.

3.3.2 System Architecture

The system architecture is show in Fig. 3.6. It considers three components: message handler (MH), request validation (RV), and fuzzy recommender system (FRS).

3.3.2.1 Message Handler

The message handler (MH) implements several functions to send recommendations available to either citizens or government agencies in different ways, for example, e-mail, SMS, voice mail, etc. The recommendation message is splitting according to the fuzzy sets by payment deadline, risk relevance, and citizen behavior. For each citizen C_i, a set of recommendation R_i will be send. Each R_i could be responded by the citizen; therefore, the ranking of accuracy of recommendations is computed by the ratio of response sent for each recommendation. Responses are received by a request validation (RV) component.

3.3.2.2 Request Validation

The RV implements a reactive autonomous agent as a real time supervisor a shown in Fig. 3.7. Its main function is to listen the environment, activate the process corresponding and increase the knowledge base. Figure 3.7 presents the inputs and output from this component.

The RV component is spitted by six inputs, three outputs, and six processes to control the states and one data base (refer to Table 3.4). Each input is considered as a sensor or listener, its waiting triggers or shift in the environment are designed to fire an action. The outputs are responsible to connect the message or order from the RV towards its target. The processes have to comply every rule or restriction until they get a successfully output requested. Finally, the database records each interaction and its results.

3.3.2.3 Fuzzy Recommender System

The fuzzy recommender system is divided by three sub components, the central data repository (CDR), the constraint-based engine (CBE), and the fuzzy recommendation engine (FRE).

Central Data Repository

Institutional data warehouses are widely applied in e-government solutions [2, 4, 29]. Therefore, Central Data Repository CDR implements an institutional data warehouse (DWH). Its main function is to integrate the information from large number of homogeneous and/or heterogeneous sources into an unique data storage.

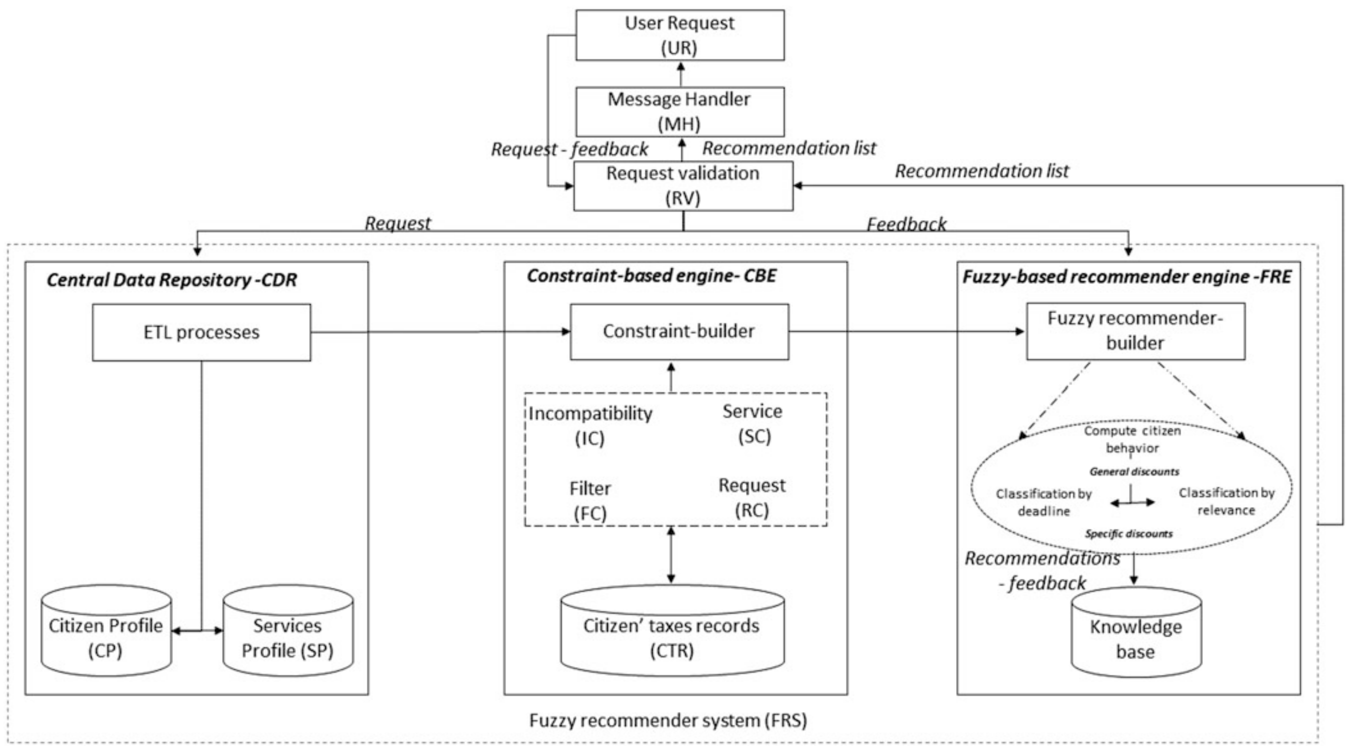

Fig. 3.6 Fuzzy-based recommender system architecture

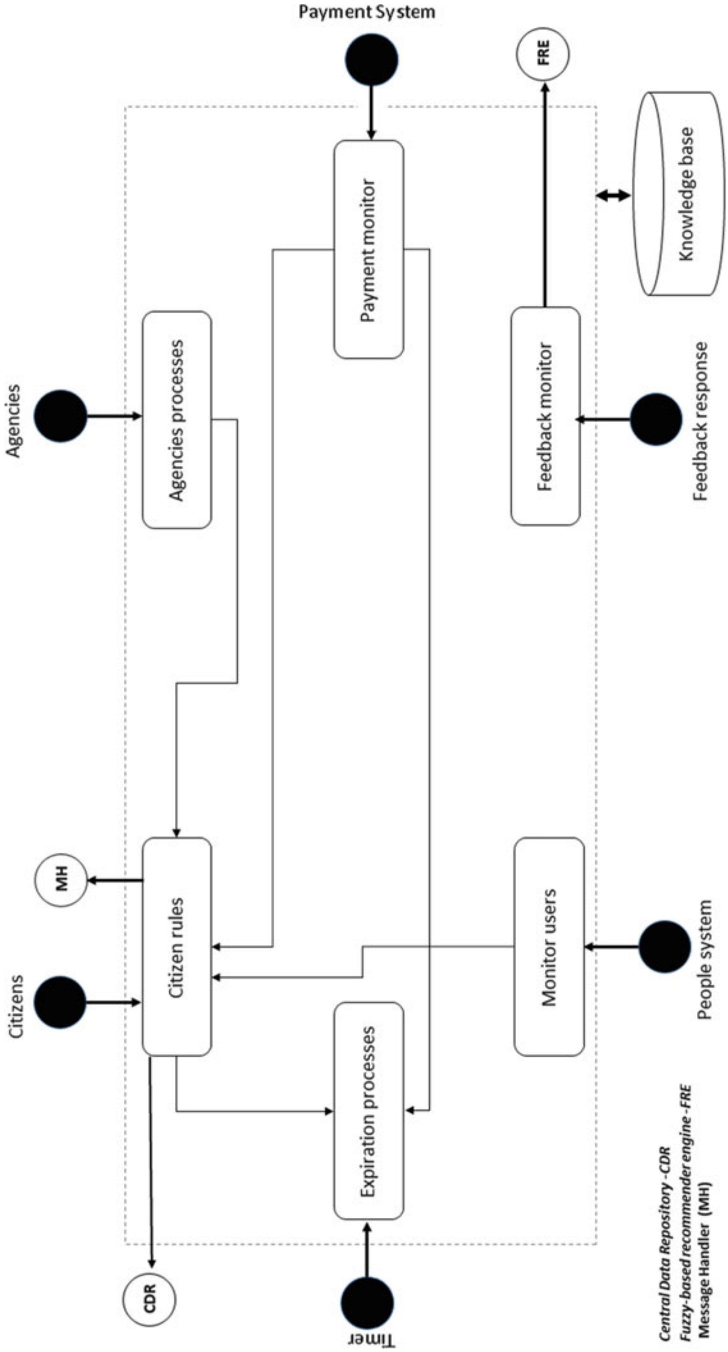

Fig. 3.7 Request validation design

Table 3.4 Elements of the request validation component

Kind	Name	Roles
Input	Citizen	It catches the actions from users registered in the website called PAM http://pam.quito.gob.ec.
Input	Agencies	It is waiting for an order from the agency operator, order such as taxes reports, recalculation of taxes, and resend recommendations
Input	Payment system	It is triggered every time that a user execute a payment by any way (e.g., credit card, banks or local cash desk)
Input	Timer	It monitors continuously the state of recommendations expiration date and throws the instructions an order to regenerate the recommendation list
Input	People system	Every time that a citizen is enrolled inside of the system it executes the recommendations related to the new citizen or data updated
Input	Feedback response	It caches every feedback from citizens related to recommendations
Output	FRE	Allows updating recommendations feedback from the citizen
Output	CDR	It starts the processes to compute the discount and recommendations
Output	MH	Entrance to recommendation lists, if needed to be sent to citizens
Processes	Citizen rules	It executes every constraints related to citizens such as notifications, change of relevance, and payment condition improved
Processes	Agencies processes	Turns on the process by specific demand
Processes	Expiration processes	It executes triggers to check the state of recommendations expiration state
Processes	Payment monitor	It connects payment services with the need to send recommendations
Processes	Monitor users	It connects payment services with need of send recommendations
Processes	Feedback monitor	It updates the recommendation with feedback into the knowledge base
Data base	Knowledge base	It records every action over recommendation and the discount generated

Moreover, the DWH has different levels of details, one of this is defined as *data mart* [24]. Data flows into the DWH from the environment which needs processes to load and transform the data, these processes are named *extraction transformation and load* (ETL). As shown in Fig. 3.7, CDR contains ETL' processes to process data from citizens and services. It's more important features are summarized in Table 3.5.

The CPD stores the profiles of involved citizens. A citizen profile (CP_i) is associated with a citizen C_i. It consists of a tuple ($CPDF_i$, $CPDP_i$, $CPDT_i$), where:

Table 3.5 Central data repository features

Name	Description	Data stored	Update mode—frequency
Citizens (CPD)	Citizen profiles and preferences	Features ($CPDF$), preferences ($CPDP$), and taxes ($CPDT$)	ETL—daily/by demand
Services (SPD)	Municipal services features	Features (SPDF), budget (SPDB), and constraints (SPDC)	ETL—by week/by demand

1. $CPDF_i$ stores the identification, social, and demographics information such as, name, date of born, gender, province, zone and so on;
2. $CPDP_i$ infers the rules from the number of times that the services (taxes) were used in the past;
3. $CPDT_i$ stores information of citizens account about pays and debts (statement of account). It includes information such as, payment date, value of pay, agency, discount.

The SPD stores the profiles of involved *eServices*. A services profile SP_i, associated with an service S_i consists of a tuple ($SPDF_i$, $SPDB_i$, $SPDC_i$), where:

1. $SPDF_i$ stores the identification, descriptions, and demographics information such as, name, date of deliver, target, zone and so on;
2. $SPDB_i$ stores the information of service budget including the cost of services without an electronic service, the cost using electronic service, and quantity citizens save using electronic services, among others;
3. $SPDC_i$ stores the information of constraints in the service such as, age, yearly income, nationality, permanent residence, physical disability, mental disability, job injury, qualification, academic degree, marital status, number of children, and relevance, among others.

The column "update mode frequency" presents the time executed by the ETL processes. Every data repository (CPD) or (SPD) are executed one time per day or according to requests made from external agents.

Constraint-Based Engine

Constraint-based Engine CBE was adapted from the constraint-based recommender systems approach in [20], and the framework proposed in [6]. Constraint-based recommender systems approach recommends services according to citizens needs [20]. In order to adapt the approach previously mentioned in the case of taxation on eGovernment services, a set of variables related to citizens profiles, services profiles, and a set of constraints are used. They define the relation between them, this set of constraints are named "Constrains-builder". Its content is presented as follows:

1. Incompatibility (IC) restricts the services according to citizens requests (e.g., age, incomes, etc.)

2. Service (SC) owns restrictions of the service (e.g., availability, date launched, budget, etc.))
3. Filter (FC) relates between citizens request and services (e.g., business taxes, urban taxes, etc.)
4. Request (RC) constraints related citizens requirements and services assigned (e.g., citizen address).

Information of relation between Citizen Profile and Service Profile in the matter of taxation is saved in Citizen Taxes Records data source (CTR). Data available on (CTR) contain information such as: Tax , citizen , payment date, value to pay, value paid.

Fuzzy-Based Recommender Engine

Fuzzy-based Recommender Engine (FRE) processes the data available on the citizen tax records data source (CTR) and prepares the recommendation list. The recommendation engine components presented in Fig. 3.6 implements the rules mentioned in Sect. 3.3.1. Recommendation lists are prepared using a set of steps defined in the fuzzy recommender builder presented as follows.

1. Citizen behaviour. It computes the citizens behavior by rating them from historical payment data. Every payment is assigned with a rating according to Table 3.2. The final rating is the average of the data set processed.
2. General Discount. With the citizens behaviour rating, the general discount percentage is computed as a membership degree by each group of fuzzy sets multiply discount range as presented in Table 3.3 and Eq. (3.3). As an example, a user rated with 3.2, his discount will be computed as membership degree of fuzzy sets *good* ($\mu_G(3.2)$) and average ($\mu_A(3.2)$) with the range of percentage settings in Table 3.3.
3. Computed classification by deadline. It computes the membership degree in the fuzzy set as shown in Fig. 3.3. Citizens with membership degree greater than 0 should be sent a paying recommendation.
4. Computed classification by relevance. It computes the membership degree in the fuzzy set as shown in Fig. 3.4. Citizens with membership degree greater than 0 should be sent a paying recommendation.
5. Computed specific discounts. It is used to compute specific discounts. The fuzzy recommender-builder implements the rules mentioned in Sect. 3.3.1.2.
6. Ranking. It represents a combination of sort processes. It implements three rankings, by discount values, risk relevance, and citizens behavior.
7. List of recommendations. The list of recommendations are performed after the classification calculation by relevance and deadline. Its responsibility is sending details by every recommendation towards the request validation (RV) component and updates the knowledge base with the new information generated.
8. Feedback. It updates the knowledge base with information responded by the recommendation sent.

Table 3.6 Land taxes restrictions

Type	Restriction	Evaluation	Approved
Incompatibility	Citizen age is higher than 18	Citizen age is 28 years old	Yes
Service	Law for land taxes in 2018 is approved	Law approved on December, 2017	Yes
Filter	Citizen is owner of apartment, house or land space	Citizen has a suite	Yes
Request	Citizen has pending invoices in order to service requested	Citizen does not pay deb in 2018	Yes

3.3.3 Recommendation Computation Process

In this section the case of the citizen of Quito city is used to follow the recommendation process. As an example, the case of the citizen Juan is used. He is 28 years old, lives in Quito in his suite, in the last 5 years he paid his municipal taxes as follows, three times with discount of 8% and two times with punish of 3%. In the present year Juan needs more information about the best way to save money. Juan makes this request on March 5, 2018 and he has a debt of USD 125.20. Juan has an account in municipal website PAM.[3] The maximum debt inside of the group of tax when Juan sent the request was USD 2125.00. The municipality has set in 2018 as max discount of 10%. From January until June, the municipality offers a discount by early payment.

According to the fuzzy-based recommender system architecture presented in Fig. 3.6, the steps to follow are:

1. Request is processed by (RV) as "Citizen" and executes "Citizen rule" process (Fig. 3.7), then a request is accepted and sent to the FRS given that Juan has an account in the municipal website.
2. Juan has data from previous year; therefore, the ETL process is skipped.
3. A set of restrictions are applied by the "Constraint-builder". Those restrictions are evaluated and the request approved as shown in Table 3.6. The data is not update in the "Citizen' taxes records", since their were not changed.
4. Juan historical payment behaviour is computed as rating of 4.2; therefore, Juan is considered as citizen between *Good and Excellent* (refer to Fig. 3.1). The rating assigned to Juan allows him to get a general discount of 37.5% (refer to Eq. (3.3)) over the max allowed by the municipality (10%). That's means a real general discount of 3.75%.

[3]PAM: http://pam.quito.gob.ec.


SOLICITUD DE IMPUESTOS - MDMQ

Inbox

servicios.ciudadanos@quito.gob.ec

to me

Apr 12

Spanish > English Translate message Turn off for: Spanish

Estimado(a) JUAN VICENTE FERNANDEZ ALCIVAR .

Estimado contribuyente, luego de haber realizado la revisión de sus impuestos prediales, se ha determinado que tiene valores pendientes de pago para el predio (335727). Sin embargo, te informamos que tu historia de pago te hace acreedor a un cliente 4 estrellas de nuestro ranking de 5. Te invitamos cordialmente acercarse a nuestros puntos de pagos y continuar mejorando tu historial crediticio. Te sugerimos las siguientes opciones de descuento para su mayor ahorro.

Descuento	Ahorro
Por relevancia de la deuda con la municipalidad.	Puedes ahorrar USD$ 8,41 con tu pago en este momento. Pagar
Por pronto pago	Puedes ahorrar USD$ 6,42 con tu pago en este momento. Pagar

QUITO ALCALDÍA

Este mensaje ha sido examinado por Symantec Messaging Gateway y se considera libre de virus y spam.

Fig. 3.8 Screenshot citizen recommendation message

5. Juan does not pay yet; therefore, he will be notified by deadline. On the other hand, he will not be notified by risk relevance, since his rating of behavior is 4.2 and his ratio of debt is 5.89%. They are not considered as relevant for early payment.
6. Juan percentage specific discounts are: deadline = 13.79%, risk relevance = 29.70%.
7. Juan will be informed with the suggested list to paying as shown in Fig. 3.8.

3.4 Model Simulation

This section evaluates the outcomes of the proposed model using a sample from the Municipality of Quito data set. The model simulation aims are:

1. To evaluate citizens historical payment behavior
2. To verifying the data set behavior with the discount model proposed
3. To summarize the outcomes over the data simulated

This section is organized as follow: in the first part the data set structure is analyzed, after that the simulation design is presented, and finally, some outcomes are presented.

3.4.1 Dataset Acquisition

In order to prepare the simulation design, the data sets from municipality taxation systems were collected, prepared, and analyzed. The dataset used has the structure presented in the Table 3.7.

Table 3.7 Data set structure for taxes payment

Name	Type	Description
CitizenID	Numeric	Unique identifier for citizen
TaxID	Numeric	Unique identifier for services or tax
TaxGroupID	String	Unique identifier for tax group, for instance, urban or business taxes
PayCompulsoryDate	Date	Deadline by type of tax
RealPayDate	Date	Real date of pay
ValueRequested	Decimal	Value requested to pay
ValuePaid	Decimal	Real income by pay
PayCondition	Char	It stores different states for citizens according to tax as follows: 1 discount, 2 no discount, 3 penalty, 4 penalty notification, 5 penalty after judicial notice
PeriodValue	Numeric	Period of tax by year or semester

Table 3.8 Municipal taxes data set

Feature	Number	Description
Total records	10,286.201	Total number of register available for processing
Citizens processed	1135.733	Citizens that have urban or business taxes
Type of taxes	20	Type of taxes available were classified by two groups: urban and business taxes
Tax groups	2	Deadline by type of tax
Period processed	34	It contains data from period 1985–2018

Table 3.9 Kind of municipal urban taxes processed

TaxId	Name
1	Urban predial
2	Urban predial parish
3	Rural predial
4	Pavements
5	Building site
34	Urban determinations
35	Rural determinations
50	Predial recalculation year 2012–2013
51	Predial recalculation year 2014

The structure shown in Table 3.7 was filled up from citizen profiles (CP) and service profiles (SP) using an ETL processes of central data repository (CDR) (refer to Fig. 3.6). Data processed is summarized in Table 3.8.

Data sets acquisition was applied to Urban Taxes, this group considered a 194.971 citizens that had to pay urban taxes in some one of the categories presented in Table 3.9.

3.4.2 Simulation Design

In order to prepare scenarios of simulation models, several steps were completed, as so also tools of support were built (refer to Fig. 3.9).

3.4.3 Outcomes

- The average rating for citizens was 3.44/5 stars. This sample was split according to pay behaviour: discount payment 9%, non discounted payment 53%, penalty payment 35%, and payment after judicial notice 3%.
- The citizens that paid with discount their mean of discount was 5.01% and their deviation 2.91; therefore, the discounts gotten by citizens were between 2% to 8%.
- The discount approach proposed presented a mean of discount of 7.68% and it deviation 0.064, therefore, the tax discount distribution is more close between citizens.

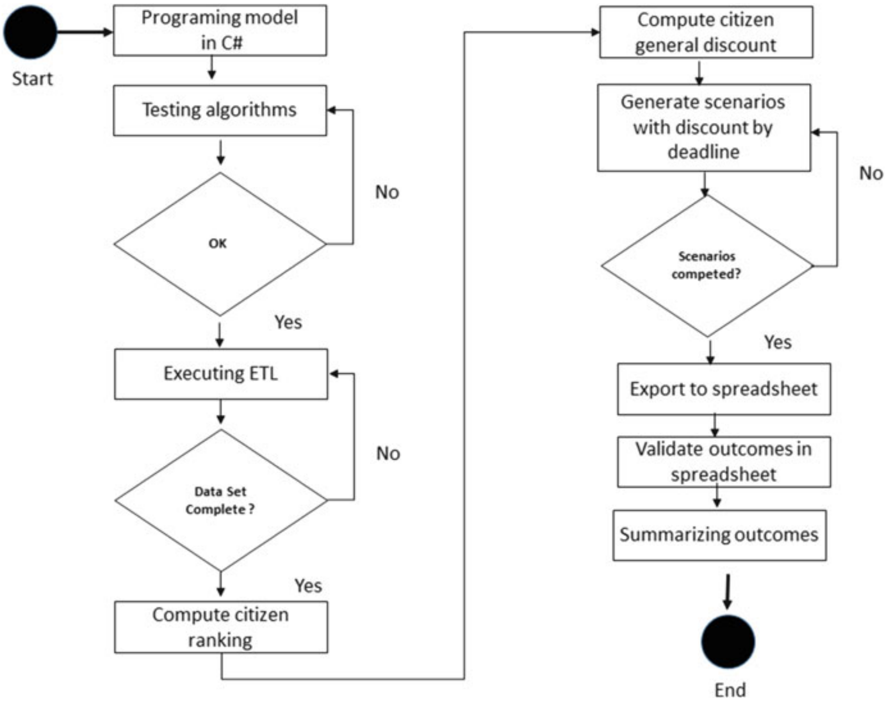

Fig. 3.9 Simulation design process

- The citizens that paid with discount in the past could get at least 4.9% of discount in future payments.
- Three scenarios were simulated to specific discount. In the worst case scenario citizens could get at least 5.98% of discount in future payments.
- The citizens that never paid with discount in the past could get at least 5.7% of discount in future payments.

Outcomes presented show evidence that the discount strategic used by the municipality was not adequate given that percentages of taxpayers that paid with discount are low. On the other hand, the discount tax model applied gives insights about its opportunities to improve municipally incomes using this strategy.

3.5 Conclusions, Limitations, and Future Work

In this work, a fuzzy-based recommender system was introduced. It was applied to notifying citizens regarding public tax payments and also discount opportunities. The proposed model has been tested using a simulation from the Municipality of Quito dataset. In this work, the model presented behaviors with a positive impact in citizens tax awareness.

This hypothesis is supported in the highlighted results of the simulations. For instance, 91% of citizens are below the level of relevance two; however, their payment behavior rating was presented inside the average. On the other hand, the rest of the citizens payment behavior ratings were consider as regular. The proposed model shows an effective way to identify and recommend this group of citizens.

The discount approach proposed in the model, show effectiveness in the way for compute taxes, therefore, that is an initial point in order to the municipality rethinking the strategies of discounts.

Another conclusion is that fuzzy logic approach applied as a classification method improves the discount balance between citizens and it gives broad options towards new discounts methods.

The recommendation approach presented in this work is centered on citizens behavior as strategy to recommender payment options to compute discounts scenarios by citizen.

Future work should focus on the design machine learning approach to improve the real time supervisor; combining both CF and reclusive approaches to propose recommendations using group preferences by finding other citizens who could have similar behaviors. The next stage for the proposed model is to test it with the citizens from the Municipality of Quito. Data gathered with citizens interactions will be used for impact evaluation and redesign of the proposed RS approach.

Acknowledgements Authors would like to thank the members of Information System Research Group (http://diuf.unifr.ch/is/) at the University of Fribourg for contributing with valuable thoughts and comments. We specially thank Prof. Dr. Andreas Meier for his support and collaboration.

References

1. Adewole, K. S., Ajiboye, A. R., Isiaka, R. M., & Babatunde, R. S. (2012). A fuzzy logic approach for evaluation of government performance in ICT projects implementation. *Journal of Emerging Trends in Computing and Information Sciences, 2*(11), 1487–1494.
2. Almutairi, B., & Khan, A. R. (2016). Persistent architecture for optimizing web service for e-government implementation. In *2016 IEEE International Symposium on Systems Engineering (ISSE)* (pp. 1–4).
3. Andino, M., Urbina, M., & Carrasco, C. (2012). *Una Nueva Política Fiscal para el Buen Vivir: La equidad como soporte del pacto fiscal*. SRI.
4. Arora, R. K., & Gupta, M. K. (2017). e-governance using data warehousing and data mining. *International Journal of Computer Applications, 169*, 28–31.
5. Asamblea Nacional del Ecuador. (2010). *Codigo organico organizacion territorial autonomia descentralizacion*. Technical report, Asamblea Nacional del Ecuador.
6. Ayachi, R., Boukhris, I., Mellouli, S., Ben Amor, N., & Elouedi, Z. (2016). Proactive and reactive e-government services recommendation. *Universal Access in the Information Society, 15*(4), 681–697.
7. Banco Central del Ecuador. (2017). Sector Fiscal. https://www.bce.fin.ec/index.php/component/k2/item/765. Accessed September 15, 2017.
8. Benhamdi, S., Babouri, A., & Chiky, R. (2017). Personalized recommender system for e-learning environment. *Education and Information Technologies, 22*(4), 1455–1477.
9. Centro de Estudios Fiscales. (2017). Presión Fiscal. https://cef.sri.gob.ec/mod/page/view.php?id=11066. Accessed September 15, 2017.
10. Contraloria General del Estado. (2017). Consultas sobre informe aprobados. http://www.contraloria.gob.ec/WFDescarga.aspx?id=48261&tipo=inf. Accessed May 22, 2018.
11. Cornelis, C., Guo, X., Lu, J., & Zhang, G. (2005). A fuzzy relational approach to event recommendation. In *Proceedings of 2nd Indian International Conference on Artificial Intelligence (IICAI)* (Vol. 5, pp. 2231–2242).
12. Cornelis, C., Lu, J., Guo, X., & Zhang, G. (2007). One-and-only item recommendation with fuzzy logic techniques. *Information Sciences, 177*(22), 4906–4921.
13. Diario -El Comercio. (2015). La recaudación de impuestos aumentó en 5 años. http://www.elcomercio.com/actualidad/quito/recaudacion-de-impuestos-aumento-anos.html. Accessed May 22, 2018.
14. Diario -El Comercio. (2017). 42 millones más por el predial en Quito en el 2018. http://www.elcomercio.com/actualidad/impuesto-predial-recaudacion-quito-municipio.htmll. Accessed May 22, 2018.
15. Diario -El Comercio. (2018). 5 277 solicitudes se procesan para cambio de avalúo predial en Quito. http://www.elcomercio.com/actualidad/solicitudes-avaluo-predios-quito-municipio.html. Accessed May 22, 2018.
16. Diario -El Comercio. (2018). Rodas: 'He dispuesto la revisión de los avalúos de todos los predios de la ciudad que sufrieron un incremento en su valor'. http://www.elcomercio.com/actualidad/quito-revision-avaluos-reclamos-predial.html. Accessed May 22, 2018.
17. Diario -La hora. (2012). Recaudación de impuestos es mayor : Noticias de Quito : La Hora Noticias de Ecuador, sus provincias y el mundo. https://lahora.com.ec/noticia/1101442386/recaudacic3b3n-de-impuestos-es-mayor. Accessed May 22, 2018.
18. Donzé, L., & Meier, A. (2012). Applying fuzzy logic and fuzzy methods to marketing. In *Fuzzy methods for customer relationship management and marketing* (pp. 1–14). Hershey: Business Science Reference.
19. European Commission. (2015). eGovernment & Digital Public Services. https://ec.europa.eu/digital-single-market/en/policies/egovernment. Accessed September 15, 2017.
20. Felfernig, A., & Burke, R. (2008). Constraint-based recommender systems: Technologies and research issues. In *Proceedings of the 10th International Conference on Electronic Commerce, ICEC '08* (pp. 3:1–3:10). New York: ACM.

21. Gräßer, F., Malberg, H., Zaunseder, S., Beckert, S., Schmitt, J., Klinik, S. A., et al. (2016). Application of recommender system methods for therapy decision support. In *IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 1–6). Piscataway: IEEE.
22. Guo, X., & Lu, J. (2007). Intelligent e-government services with personalized recommendation techniques. *International Journal of Intelligent Systems, 22*(5), 401–417.
23. Hsu, T.-H., Chu, K.-M., & Chan, H.-C. (2000). The fuzzy clustering on market segment. In *The Ninth IEEE International Conference on Fuzzy Systems, FUZZ IEEE 2000* (Vol. 2, pp. 621–626). Piscataway: IEEE.
24. Inmon, W. H. (2002). *Building the data warehouse*. Hoboken: Wiley.
25. Jones, E. E., & Chikezi, D. (2016). Assessment of the impact of tax reforms on economic growth in Nigeria. *Journal of Accounting and Financial Management, 2*(2). ISSN 2504-8856.
26. Lu, J., Shambour, Q., Xu, Y., Lin, Q., Zhang, G. (2013). A web-based personalized business partner recommendation system using fuzzy semantic techniques. *Computational Intelligence, 29*(1), 37–69.
27. Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G. (2015). Recommender system application developments: a survey. *Decision Support Systems, 74*, 12–32.
28. Meier, A., & Stormer, H. (2009). *eBusiness & eCommerce: Managing the digital value chain*. New York: Springer.
29. Molnár, B. (2017). Proposal for application of data science methods in e-government. In A. Kő & E. Francesconi (Eds.), *Electronic government and the information systems perspective* (pp. 143–157). Cham: Springer.
30. OECD. (1996). Definition of Taxes. http://www.oecd.org/daf/mai/pdf/eg2/eg2963e.pdf. Accessed September 15, 2017.
31. Paz y Miño, J. (2015). *Historia de los Impuestos en Ecuador*. Quito: SRI-PUCE.
32. Porcel, C., Ching-López, A., Tejeda-Lorente, A., Bernabé-Moreno, J., & Herrera-Viedma, E. (2018). Analysis of different proposals to improve the dissemination of information in university digital libraries. In J. Kacprzyk, E. Szmidt, S. Zadrożny, K. T. Atanassov, & M. Krawczak (Eds.), *Advances in fuzzy logic and technology 2017* (pp. 195–206). Cham: Springer.
33. Quito Informa. (2018) Sistema para revalúo de Impuesto predial está listo y operativo. http://www.quitoinforma.gob.ec/2018/02/19/sistema-para-revaluo-de-impuesto-predial-esta-listo-y-operativo/. Accessed May 22, 2018.
34. Rist, T., Seiderer, A., Hammer, S., Mayr, M., & André, E. (2015). Care - extending a digital picture frame with a recommender mode to enhance well-being of elderly people. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)* (pp. 112–120).
35. Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. In *Applications of data mining to electronic commerce* (pp. 115–153). Berlin: Springer.
36. Servicio de Rentas Internas. (2015). Derechos del Contribuyente. http://www.sri.gob.ec/web/guest/derechos-del-contribuyente. Accessed September 15, 2017.
37. Tamayo, L. F. T. (2014). *SmartParticipation*. Berlin: Springer.
38. Terán, L., & Meier, A. (2010). A fuzzy recommender system for eElections. In *EGOVIS* (pp. 62–76). Berlin: Springer.
39. Werro, N. (2015). *Fuzzy classification of online customers*. Berlin: Springer.
40. Wu, D., Lu, J., & Zhang, G. (2015). A fuzzy tree matching-based personalized e-learning recommender system. *IEEE Transactions on Fuzzy Systems, 23*(6), 2412–2426.
41. Yager, R. R. (2003). Fuzzy logic methods in recommender systems. *Fuzzy Sets and Systems, 136*(2), 133–149.
42. Yager, R. R., & Yager, R. L. (2012). Fuzzy soft social network modeling and marketing. In *Fuzzy methods for customer relationship management and marketing: Applications and classifications* (pp. 16–40). Hershey: Business Science Reference.
43. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353.

44. Zadeh, L. A. (1983). A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with Applications, 9*(1), 149–184.
45. Zhang, S., Xi, C., Wang, Y., Zhang, W., & Chen, Y. (2013). A new method for e-government procurement using collaborative filtering and Bayesian approach. *The Scientific World Journal, 2013*, Article ID 129123, 10pp. https://doi.org/10.1155/2013/129123