SmartParticipation – A Fuzzy-Based Platform for Stimulating Citizens' Participation

Luis Terán
University of Fribourg

Andreas Meier
University of Fribourg

Abstract

Social networks and communities have become an important environment for exchanging information about products, services, music, and movies, among other things. In an information and knowledge society, such technologies could also improve democratic processes, increase citizens' interest in political issues, enhance participation, and renew civic engagement. However, the difficulty of finding other citizens or parties that share common goals is still a barrier. In this paper, a fuzzy-based recommender system architecture for stimulating political participation and collaboration is proposed. As a first step, citizens and candidates must create their profile using a fuzzy interface to answer political questions. Afterwards, the recommendation engine computes fuzzy clusters of politicians. Finally, citizens can evaluate the top-n politicians from different political parties in an election process and/or build up communities that share their common interests. The SmartParticipation project uses the database of smartvote, a well-known voting advice application (VAA) for local, cantonal, and national elections in Switzerland. The recommendation engine works with a modified fuzzy c-means algorithm and the Sammon mapping technique used for visualizing recommendations. First results demonstrate the potential for building political communities and the stimulation of civic participation.

Keywords

Recommender System, Community Building, eParticipation, eDemocracy, eCommunity, Fuzzy Clustering.

1. Introduction

In recent years, governments have taken promising steps toward deploying electronic services to citizens and companies. The European Union, for instance, is beginning the transition to a knowledge-based society¹ by implementing Web-based services for citizens, such as income taxes, job search services, social security contributions, certificates, public libraries and schools, and e-health, among others. In addition, public services for business are offered: social contributions for employees, declarations of taxes, new company registration, customs declaration, and public procurement, among others.

Strengthening democratic decision making and Web-based participation still remains a challenge. Political decision processes have become more complex, involving a wide range of stakeholders and interests and demanding more difficult

trade-offs. Using Web technologies that facilitate transparency and public involvement could stimulate *eDemocracy* and participation.

This research combines Web-based voting advice applications (VAAs) with fuzzy logic. With fuzzy classification, a politician or citizen can belong to more than one class with differing degrees of membership. This notion of membership not only provides a better description of political parties but also helps users make differentiated decisions in *eElection* processes (Terán and Meier [11], Terán et al. [10]) and when building political communities.

The fuzzy-based recommendation engine used in the *SmartParticipation* project for citizens and politicians uses data from the *smartvote* project². Since 2003, national, cantonal, and local elections in Switzerland have used the *smartvote* approach for *eElections*. Voters and politicians specify their political profile by answering the same set of questions. The *smartvote* web site is a political issue-matching system with analytical tools for evaluating and analyzing political positions (Fivaz and Schwarz [3]).

In the fuzzy-based prototype, a modified fuzzy c-means algorithm and the *Sammon mapping* technique are used for clustering politicians and citizens. Applicants can therefore use the fuzzy recommender systems to evaluate the top-n neighbors (politicians and/or citizens) with a fuzzy similarity range, clusters of politicians and citizens, or build up new political communities.

This research paper is structured as follows: Section 2 describes the eGovernment Framework used, and the SmartParticipation project. Section 3 presents a description of so-called online VAA and their influence on voters' decision making. Section 4, delineates the fuzzy-based recommender system architecture. It consists of an overview and the description of building blocks, such as user profile generation, the recommender engine, fuzzy clustering, the top-n recommendations, and political community building. Section 5 presents the prototype and first results using the *smartvote* dataset (the national Swiss elections of 2007). Section 6 proffers concluding remarks and suggestions for future research.

2. eGovernment Framework

The eGovernment Framework of the University of Fribourg, which is used by the SmartParticipation project, is a process-oriented maturity model with three important levels (illustrated in Fig. 1 below).

¹See: http://ec.europa.eu/information_society

²*smartvote* project: <http://smartvote.ch>

Figure 1: eGovernment Framework of the University of Fribourg, adopted from Meier, p. 7 [6].

Process Level I - Information and Communication.

The lowest level provides information and communication facilities for *eGovernment* (*eAssistance*). It focuses on the design of communal Web portals, respectively more extensive eGovernment portals, and the use of Web 2.0 technologies. For barrier-free Web access, compliance with the Accessibility Guidelines (WCAG) proposed by the World Wide Web Consortium (W3C) is imperative.

Process Level II - Production. The middle process consists of administrative public services required for options A2A (Administration to Administration), A2C (Administration to Citizens), and A2B (Administration to Business). These services include management services for electronic procurement (*eProcurement*); traditional services, such as taxation, residents' registration, identification acquisition, e-health, and public education (*eServices*); contracts based on digitally signed electronic documents (*eContracting*) and fulfillment elements, such as electronic shipment, electronic payment, and the assurance of data security and safety (*eSettlement*).

Process Level III - Participation. Notably, the top level emphasizes civic collaboration and participation, requisites of a progressive and responsible knowledge society. Virtual forms of collaboration (*eCollaboration*), including the utilization of social and semantic Web technologies, enable the knowledge society to develop further. In addition to electronic voting (eDemocracy), communication platforms, including processes to build new communities and political networks (*eCommunity*), stimulate participation and knowledge exchange between citizens.

Evaluating public services for eGovernment, most countries are focusing on information exchange and support for administrative processes in public affairs. In other words, governmental authorities have implemented process elements of the two lower levels (I and II) in the eGovernment Framework on Fig. 1.

2.1 eDemocracy

The term eDemocracy refers to the use of information and communication technologies to enable citizens to exercise their rights and fulfill their obligations in the information and knowledge society in a time- and place-independent manner.

In his work, Meier[6] mentions the importance of citizen participation in eDemocracy (e.g., eElection and *eVoting*). Meier defines the term *eDiscussion* as a stage in which citizens could know more about the candidates or the subject during the voting process. eDiscussion uses information and communication technologies, such as discussion forums, decision making aids, and subscription services, to aid voters (users) in making decisions.

According to Meier, the next stage of eDemocracy, following eVoting and eElections, is *ePosting* as another stage that is required on eDemocracy. This stage facilitates the publication of results and gives voters (users) the chance to open up discussion channels about eVoting and eElection.

2.1.1 Process steps for eVoting and eElection

Electronic votes and elections differ from traditional voting and election procedures mainly in their succeeding and post-processing phases (cf. Fig. 2), if the advantages of electronic exchange relations are exploited. Through changed and expanded information and discussion of politics in the process steps eDiscussion and ePosting, it is hoped that citizens will become more involved with political issues and engage in further community building. Fig. 2 distinguishes the following process steps:

eDiscussion. Prior to a vote or election, citizens can enhance their own opinion-forming process by requesting not only information, but also opinions and evaluations from discussion forums. Furthermore, subscription services allow the citizens to draw on documents or bases of decision making, and learn about changes in and extensions of topical issues.

eVoting. Within the timespan established by the authorities, citizens can fill out their electronic ballot and submit it. Before that, they identify themselves and register with a governmental institution; the vote afterward is made anonymously. The governmental institution can add an optional survey questionnaire to the ballot, in order to, for example, get citizens' feedback on questions of procedure and implementation.

eElection. The published *spiderweb* profiles of political candidates and additional information on their abilities and skills make it easier for voters to fill out the electronic ballot (during eDiscussion step). Again, citizens must register by means of an election and checking card and request a valid ballot from the governmental institution before voting electronically. Requesting that voters answer additional optional questions may be beneficial.

ePosting. Publication of voting and election results on the eGovernment portal for the associated governmental institution is directed at citizens but can also be studied and used by organizations and the press. To that end, suitable visualization and analysis tools can be offered so that electoral and voting behavior, and voting and election results, may be analyzed and discussed. Public blogs make it possible to comment on electronic votes and elections even after election day, enabling citizens to explore relevant the topics more deeply. Apart from actual voting and election results, voting and election cards and un-ambiguous identification numbers for them should be published. By this means, all

Figure 2: eVoting and eElection as part of a process chain

citizens will be able to verify whether their vote has actually been registered and processed. This method is more transparent than traditional votes and elections and will thus help to win the citizens' trust in eVoting and eElections.

The description of the process steps eDiscussion, eVoting, eElection and ePosting shows that the use of electronic information and exchange relations increases citizens' involvement and stimulates public discussion.

2.2 eCommunity

The Internet is developing into an environment in which citizens display themselves, meet with others, exchange information and services, promote common projects, and overcome linguistic and cultural boundaries. This section presents some alternative communication and Web-based tools for community formation, demonstrated in Meier's work[6].

Computers and communication channels not only serve collaborators in the administration, enabling them to handle their workload, but also make encounters and communities possible. In the same way that street cafes, markets, and exhibitions serve as points of encounter in real life, besides home and the work place, the computer network of networks develops into a virtual location. Topic-specific, cultural, or scientific meeting points on the Internet engender a new kind of community building.

Computer networks are populated by citizens and avatars³. The Internet, or cyberspace, can enhance one's living environment. As in real living environments, infrastructures are developed for virtual space; platforms for exchange are supplied and services offered. In addition, rules of conduct and protective measures are implemented with the aim of safeguarding privacy and preventing misuses.

Among the communities created on the Internet, we can distinguish between two kinds:

Communities of interest. These comprise citizens who share interest in a common thing or share a hobby.

Communities of practice. These comprise groups of citizens who participate in a project together for a governmental institution, investing time and knowledge.

Both kinds of communities can benefit from information and communication systems. Community support systems serve as a meeting place for members, as well as a place for them to exchange know-how and master tasks or challenges.

³In Hinduism, avatars are the embodiment of a god on earth; and in today's information society, they are images representing people who adopt a fictional identity on the Internet.

Web-based platforms not only facilitate the development of communities, but also make it possible for people to meet other community members and utilize the community's collective know-how.

Web-based platforms and the corresponding software systems for community formation can be characterized as follows:

Civic Network Systems. Civic networks, or community networks, are electronic meeting points for citizens whose common ground is a shared place or living environment. This may be a city or a mountain area, whose inhabitants want to meet with their cohabitants in virtual space. Apart from designing and moderating discussion forums, with civic networks systems, the focus lies on community projects or training and education programs.

Buddy Systems. Buddy systems show where colleagues or friends are currently located and how they can be reached electronically. Group members' social or task-oriented perceptions (awareness) makes it possible to meet virtually or exchange experiences; apart from that, such systems indicate when participants do not want to be disturbed. For actual reunions, the systems can establish audio or video connections to overcome spatial distance (media space). Recent developments have allowed for the creation of three-dimensional worlds and replace real areas or points of encounter with virtual ones (virtual realities).

Matchmaking Systems. The term matchmaking originally referred to marriage bureaus, but it is conceived as a wider concept in cyberspace. Matchmaking is about establishing relations of social and economic exchange relations. These systems promote contacts and activities in a commonly used environment. For example, networks of acquaintances are utilized to make new contacts, based on an already existing mutual trust, and exchange information.

Recommender Systems. These systems are about learning Internet users' preferences and making suggestions for their further development. Special procedures (collaborative filtering) make it possible to categorize the participants' preferences and pass on suggestions for activities and further education that are important for a certain group. For example, if someone wishes to become acquainted with a new subject area, such systems recommend suitable literature, possibly complemented by an expert evaluation.

Corporate Blog Systems. Corporate Weblogs are up-to-date digital journals created by groups of people or organizational units. These systems or platforms make it possible to mediate knowledge, discuss topics, and nurture relationships. Corporate blogs serve to support organizational goals and can usually be subscribed to by all interest groups. To further elucidate the topic, the applications of corporate blogs will be discussed, as they assume different functions in an organization.

2.3 The SmartParticipation Project

The SmartParticipation research project is used in the top level of the eGovernment Framework (refer to Fig. 1). It provides citizens with a simple and innovative alternative

based on fuzzy clusters for monitoring and evaluating the performance and objectives of political actors. Additionally, this application allows citizens to create virtual communities based on their profiles, such as new political parties, thematic groups and civic networks, participating in national issues by opening channels of discussion and debate through the use of information and communication technologies (ICTs) and Web 2.0.

Two main tools have been designed in the SmartParticipation application: fuzzy clustering and Virtual Communities. Fig. 3 illustrates the use of such tools on different topics related to eCollaboration, eDemocracy and eCommunity. Further sections describe in greater detail their use.

Figure 3: The SmartParticipation project.

The current recommendation engine, described in this paper, has been designed for eDemocracy and eCommunity processes. Further work will also cover eCollaboration.

3. Voting Advice Applications

The amount of data available on the Internet is growing rapidly, a phenomenon that affects not only our daily lives but it also politics and electoral campaigns.

VAA's are Web-based systems, which provide voters with information about that political party, or candidate is closest to their preferences and political values. Voters are asked to create a political profile by filling out a questionnaire on different political issues. Then the VAA compares their answers with the positions of parties or candidates in the system who have also completed the questionnaire. Finally, voters are provided with a voting recommendation in the form of a list ranking parties or candidates according to the degree of their issue congruence with the particular voter.

VAA's⁴ are quite diverse. They vary in design as well as the features they offer, but in the end, they all share the same key functions. According to Ladner et al.[5], the first operational VAA was the Dutch StemWijzer⁵. It went online for the first time in 1998 and provided 250,000 people with

⁴Examples of VAA's are: smartvote (smartvote.ch), Who do I vote for? (www.whodoivote.co.uk), Project Vote Smart (www.vote-smart.org), Glassboth (www.glassboth.org), On the Issues (www.ondtheissues.org), Openpolis (www.openpolis.it), smartvote Ottawa (www.smartvotetottawa.ca), Political Compass (www.politicalcompass.org), and EU-Profiler (www.euprofiler.eu).

⁵StemWijzer: (www.stemwijzer.nl)

voting advices. In 2006, this figure exploded to 4.7 million voting advices, which represents 40% of the Dutch electorate (Walgrave et al.[13]). Fivaz et al's.[2] work proffers clear evidence of the increasing popularity of VAA's.

3.1 Impacts of VAA's on Voters' Decision Making – The Smartvote Case

After the 2007 national elections, a survey among smartvote users was conducted. Among other aspects, this survey data allows a look at the impact smartvote has had on voters and their decision making process. The initial analysis of this data showed that its users regarded smartvote as an important information channel (Fivaz and Felder[2]). In actuality, they regarded smartvote as their most important information source.

Furthermore, Ladner et al.[5] found evidence, based on the same survey data, that smartvote also directly affected voters' decisions: 67% of the smartvote users stated that the voting recommendation that they received influenced their electoral choice. This finding is supported by additional evidence. Ladner et al.[5] showed that among smartvote users, the number of swing-voters, or those who voted for a different party than in the previous election, is significantly higher than among voters who did not use smartvote. However, since these are only initial results, they should be treated with caution. Nevertheless, they clearly indicate that the smartvote voting recommendations do affect electoral decisions. Thus, researchers must examine the applied methods for calculating smartvote's data more closely.

Schwarz et al's.[9] analysis of congruence in the Swiss lower house between 2003 and 2009 indicates that 85% of elected authorities voted in the parliament according to what they claimed on smartvote's questionnaire when they were candidates.

The recommendation system presented in this work could be used with the approach proposed by Schwarz et al.[9], in which 34 smartvote's questions came up in the parliament. This approach allows one to evaluate whether what candidates claimed before elections was reflected in their actions as elected authorities. For that reason, if analysis of congruence is required the design of the questions used to generate candidates' profiles is extremely important.

4. Fuzzy-Based Recommendation Engine

Although collaborative filtering-based approaches are more widely used for recommender systems in *eCommerce* to suggest items that a customer is presumably going to buy, they are only suitable in the repeat-appeared scenario, which is described by Vozalis and Margaritis[12].

Recommender systems for eDemocracy must also be suitable in the one-and-only items scenario, introduced by Guo and Lu[4], in which the recommendation target is a unique item/event (e.g., a voter v wants to receive a recommendation of n candidates that are close to his preferences in an election E).

In [14], Yager makes a distinction between "recommender systems" and "targeted marketing." Yager considers recommender systems to be "participatory" systems, where users intentionally provide information about their preferences. In contrast, targeted marketing methods are based on extensional information, which refers to actions or past experiences with specific objects.

The definition of recommender systems as introduced by Yager is used in this paper. In addition, it is assumed that users are willing to provide information about their preferences.

In the voting/election and the community-building scenarios, the recommendation makes no use of past events, given the fact that candidates/citizens could be different for each election/recommendation, or they could change their political orientation.

Furthermore, during recommendation generation, it is necessary to define the elements needed and the output of the system that is developed. As mentioned previously, a recommender system that is used for eDemocracy and eCommunity must be able to provide a list of n candidates/citizens who most closely match the preferences of a specific voter (user).

A fuzzy-based cluster algorithm for recommendations in eElections, originally introduced by Terán and Meier [11], is extended for the creation of political communities. It provides information about the closest candidates/citizens to a voter (user), and the distribution of political parties organized in fuzzy clusters forming political communities.

4.1 Fuzzy Logic, Fuzzy Sets, and Fuzzy Clustering

The SmartParticipation project uses a fuzzy-based approach, for the creation of political/thematic groups, assuming that citizens/candidates can not be considered unique items. People's political orientation can evolve or change over time. Therefore, using sharp clustering does not make sense even in the case of politicians who in most cases are part of a specific political party. According to their personal profiles, they may stand at different distances from the centers of political parties. The results shown in this paper in section 5 using the SmartParticipation prototype demonstrate the assumptions made.

This section introduces the basic concepts of fuzzy logic, fuzzy sets, and fuzzy clustering are introduced.

Fuzzy Logic. Fuzzy Logic is a multi-value logic that allows a better understanding of the result of a statement more approximate than precise in real life. In contrast with "sharp logic," where the results of a statement are binary ("true or false" or "one or zero"), fuzzy logic admits a set of truth-values in the interval $[0, 1]$. Fuzzy logic is derived from fuzzy set theory, introduced by Zadeh in [15], where a fuzzy set is determined by a membership function with a range of values between 0 and 1. In [15], Zadeh gives a definition of fuzzy sets, which is shown below:

DEFINITION 1. A fuzzy set is built from a reference set called a universe of discourse. The reference set is never fuzzy. Assuming $U = \{x_1, x_2, \dots, x_n\}$ as the universe of discourse, then a fuzzy set A in U ($A \subset U$) is defined as a set of ordered pairs:

$$\{(x_i, \mu_A(x_i))\}$$

where $x_i \in U$, $\mu_A : U \rightarrow [0, 1]$ is the membership function of A , and $\mu_A(x) \in [0, 1]$ is the degree of membership of x in A .

Fuzzy Clustering. Fuzzy clustering is an unsupervised learning task that aims to decompose a set of objects into

"clusters" based on similarities, where the objects belonging to the same cluster are as similar as possible. In contrast, using sharp clustering, each element is associated with just one cluster.

The main algorithms used to generate clusters are c-means (sharp clustering) and fuzzy c-means (fuzzy clustering). The fuzzy recommender system proposed in this paper uses the fuzzy c-means algorithm explained below.

The c-means algorithm, originally proposed by Bezdek [1], is a method of cluster analysis that aims to partition n observations into c clusters. Each observation belongs to one and only one cluster. The recommendation approach presented in this paper uses the fuzzy c-means algorithm, which is described in more detail as follows.

Fuzzy C-Means Algorithm. The fuzzy c-means algorithm (FCM) is an extension of the c-means algorithm that allows for gradual membership of data points to clusters with different degrees of membership according to the fuzzy set theory introduced by Zadeh [15]. Thus, fuzzy c-means defines a given set of samples $X = \{x_1, \dots, x_n\}$, a set of clusters Γ_i ($i = 1, \dots, c$ and $2 \leq c < n$), and a $c \times n$ fuzzy partition matrix $\vec{U} = [u_{ij}]$. The membership degree u_{ij} of an observation x_i in a cluster Γ_i is such that:

$$u_{ij} = \mu_{\Gamma_i}(x_j) \in [0, 1]$$

A probabilistic cluster partition defined by the constraints in (1) and (2) guarantee that clusters are not empty, and that the sum of the membership for each x is equal to 1

$$\sum_{j=1}^n u_{ij} > 0, \forall i \in \{1, \dots, c\} \quad (1)$$

$$\sum_{i=1}^c u_{ij} = 1, \forall j \in \{1, \dots, n\} \quad (2)$$

Thus, the FCM algorithm is based on minimization of the objective function shown in (3)

$$J_m = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - y_i\|^2 \quad (3)$$

where x_j is the j -th element of d -dimensional measured data, y_i is the d -dimensional center of cluster i , m is any real number greater than 1 (m determines the level of "fuzziness"; $m = 2$ is a typical value used), and $\|*\|$ is any norm expressing the similarity between any measured data and the center. In (3), $\vec{Y} = [y_i]$ is the matrix of cluster centers ($i = \{1, \dots, c\}$).

The membership function u_{ij} and the center of clusters y_i are computed by taking the derivative of the objective function J_m with respect to the parameters to optimize equal to zero. Taking into account constraint (2), equations (4) and (5) are obtained.

$$u_{ij} = \frac{1}{\sum_{l=1}^c \frac{\|x_j - y_l\|^2}{\|x_j - y_i\|^2}^{m-1}} \quad (4)$$

$$y_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (5)$$

The FCM algorithm is a two-step iterative process defined as follows. First, set the input variables c , m , and ϵ (ϵ is a termination criterion, normally $\epsilon \in [0, 1]$). Second, set an iteration number $k = 0$. Third, randomly generate a matrix of cluster centers $\vec{Y}^{(k)}$. Then, given the initial matrix $\vec{Y}^{(k)}$, compute the fuzzy partition matrix $\vec{U}^{(k)}$.

Finally, using a repeat-until loop, update $\vec{Y}^{(k+1)}$ using $\vec{U}^{(k)}$ and then, update $\vec{U}^{(k+1)}$ using $\vec{Y}^{(k+1)}$. Repeat this process until the termination criterion is reached ($|\vec{U}^{(k+1)} - \vec{U}^{(k)}| \leq \epsilon$). FCM algorithm is presented in Algorithm 1.

Algorithm 1 FCM

Input: c, m, ϵ

Output: $\vec{U}^{(k+1)}$

- 1: Set iteration number: $k \leftarrow 0$
 - 2: Generate matrix of cluster centers: $\vec{Y}^{(k)} \leftarrow \text{random}$
 - 3: Compute $\vec{U}^{(k)} \leftarrow \vec{Y}^{(k)}$
 - 4: **repeat**
 - 5: Update $\vec{Y}^{(k+1)} \leftarrow \vec{U}^{(k)}$
 - 6: Update $\vec{U}^{(k+1)} \leftarrow \vec{Y}^{(k+1)}$
 - 7: **until** $|\vec{U}^{(k+1)} - \vec{U}^{(k)}| \leq \epsilon$
 - 8: **return** $\vec{U}^{(k+1)}$
-

The approach proposed in this paper for advising citizens on elections and the creation of political communities is based on the generation of a fuzzy cluster and uses a modified version of the fuzzy c-means algorithm. It is explained in more detail in section 4.5.

4.2 Architecture Overview

The recommendation procedure is divided into three steps: In the first step, the voters (users) and candidates must create their profiles using a fuzzy interface, which is a convenient tool used to determine the level of agreement, disagreement, and relevance for each specific question. The fuzzy profiles are stored in a database.

In the second step, once all necessary profiles have been created, the user selects the recommendation target and the type of output (top-N recommendation, fuzzy cluster analysis or political community).

In the final step, once the recommendation engine has computed all the information, the user receives the recommendation in the pre-established format.

The architecture of the fuzzy recommendation approach is presented in Fig. 4. Each element is presented in more detail in the following sections.

4.3 User Profile Generation

In order to provide a recommendation, voters (users) and candidates must generate a profile that describes their preferences using a fuzzy interface to complete a questionnaire regarding political issues (each question has different possible responses).

The fuzzy interface is a convenient tool used to determine the level of agreement/disagreement and relevance for a specific question. Unlike other similar tools, it provides a higher number of possibilities for each citizen/candidate to answer the questions. The interface is designed to be as intuitive and convenient as possible for users.

In addition to the fuzzy interface, the system contains a profile representation, so-called fuzzy profile (FP), which is

Figure 4: Fuzzy-based recommender system architecture

a multi-dimensional Euclidean space defined by:

$$FP_i = (fpc_{i1}, \dots, fpc_{in})$$

where FP_i is the FP vector of user i , and fpc_{ij} is the j -th fuzzy profile component (fpc). Each fpc is the norm of a multi-dimensional Euclidean space defined by:

$$fpc_{ij} = \|(q_{ij1}, \dots, q_{ijl})\| = \sqrt{\sum_{k=1}^l q_{ijk}^2}$$

where fpc_{ij} is the j -th fpc of FP_i , and q_{ijk} is the k -th component of fpc_{ij} . In order to illustrate the use of a FP, a profile instance of user i (FP_i) composed by n questions ($FP_i = (fpc_{i1}, \dots, fpc_{in})$) is used. Each question has two components: “tendency” and “relevance” (q_{ij1} and q_{ij2}), where:

$$fpc_{ij} = \|(q_{ij1}, q_{ij2})\| = \sqrt{q_{ij1}^2 + q_{ij2}^2}$$

The fuzzy interface and FP are described in greater detail in the work of Terán and Meier [11].

4.4 Recommendation Engine

The recommendation engine is based on the generation of fuzzy clusters using the modified fuzzy c-means algorithm described in greater detail in section 4.5.

Once the profiles are generated, the next step is to ask for a recommendation. At this point, the user selects a particular event and the type of recommendation (top-N recommendation, fuzzy clustering analysis or political community). The request is sent to the recommendation engine, which processes the query.

To provide a graphical representation of results that users can easily analyze, the recommendation engine transforms the high-dimensional space of profiles to a lower dimensional space (bi-dimensional), which reduces the complexity of data analysis. The recommender engine uses a mapping method originally proposed by Sammon[8] that attempts to preserve inter-pattern distances. The modified fuzzy c-means algorithm and the Sammon mapping technique are described in the following sections.

Sammon Mapping. The Sammon mapping technique is a well-know technique that transforms a high-dimensional space (n -dimensions) to a space with lower dimensionality (q -dimensions), finding N points in the lower dimensional space.

Denoting the distances between two different points x_i and x_j ($i \neq j$) in the original space as d_{ij} , and the distance between points y_i and y_j in the mapped space as d'_{ij} , then the mapping becomes a problem of minimizing Sammon's stress E , defined in (6):

$$E = \frac{1}{\lambda} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{(d_{ij} - d'_{ij})^2}{d_{ij}} \quad (6)$$

where, $\lambda = \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij}$.

In order to minimize E , Sammon applied a steepest descent technique, in which the new y_{il} at iteration $t+1$ is given by

$$y_{il}(t+1) = y_{il}(t) - \alpha \left[\frac{\frac{\partial E(t)}{\partial y_{il}(t)}}{\frac{\partial^2 E(t)}{\partial^2 y_{il}(t)}} \right] \quad (7)$$

where, $y_{il}(t)$ is the l -th coordinate of point y_i in the mapped space and α is a constant that has been computed empirically to be $\alpha \approx 0.3$ or 0.4 . The partial derivatives in (7) are given by

$$\begin{aligned} \frac{\partial E(t)}{\partial y_{il}(t)} &= -\frac{2}{\lambda} \sum_{k=1, k \neq i}^N \left[\frac{d_{ki} - d'_{ki}}{d_{ki} d'_{ki}} \right] (y_{il} - y_{kl}) \\ \frac{\partial^2 E(t)}{\partial^2 y_{il}(t)} &= -\frac{2}{\lambda} \sum_{k=1, k \neq i}^N \frac{1}{d_{ki} d'_{ki}} \\ &\quad \cdot \left[(d_{ki} - d'_{ki}) - \left(\frac{(y_{il} - y_{kl})^2}{d'_{ki}} \right) \left(1 + \frac{d_{ki} - d'_{ki}}{d_{ki}} \right) \right] \end{aligned}$$

This paper focuses on the Sammon mapping method for the visualization of clustering results, which preserves inter-pattern distances. The three problems that must be taken into account when the Sammon mapping technique is used, which are:

- The prototypes of clusters are usually not known a priori. They are generally calculated while partitioning of the data. These prototypes can be vectors that are dimensionally equal to the examined data points, but they can also be defined as geometrical objects (i.e., linear or non-linear sub-spaces or functions). Sammon mapping is a projection method that is based on the preservation of the Euclidian inter-point distance norm, so it can only be used by clustering algorithms that are calculated with this type of distance norm. As mentioned in section 4.3, FPs are defined to be a multi-dimensional Euclidean space, which fulfills the required condition of the Sammon mapping technique.
- The Sammon mapping algorithm forces one to find, in a high n -dimensional space, N points in a lower q -dimensional subspace, such these inter-point distances correspond to the distances measured in the n -dimensional space. This causes a computationally expensive algorithm, since every iteration step requires the computation of $N(N-1)/2$ distances.
- Finally, this gradient-descent method has the possibility of reaching a local minimum in the error surface, while searching for the minimum of E , so experiments

with different random initializations are necessary. In order to avoid this problem, the initialization is estimated using the principal component analysis (PCA) technique introduced by Pearson[7], which maps the data points into a lower dimensional space.

4.5 Fuzzy Cluster Algorithm

Once the profiles are mapped to a low-dimensional space using the Sammon mapping technique, the recommendation engine generates fuzzy clusters by using a modified FCM algorithm, which requires two main inputs: the number of clusters, and a matrix of cluster centers. For this reason, prior knowledge of the dataset is required. In the election scenario, the recommendation engine considers the number of clusters to be equal to the number of political parties.

The second input is the matrix of initial centers. In the case of FCM, is generated randomly. Consequently, the algorithm may converge to a local minimum, given its random nature. Terán and Meier[11] modified the FCM algorithm to avoid this problem. The modified algorithm initializes the matrix of centers with a random member of each political party. However, Terán et al.[10] show that in some cases, the members of a political party could belong to a different cluster from the one that corresponds to his/her own political party. To avoid this problem, the matrix of centers used in this work is initialized taking the mean average of answers from all candidates in the same political party. The initialization process is based on two assumptions: First, the cluster formation relies on the existence of political parties. Second, the members of political parties have the same ideology, according to the ACE project⁶.

The modified fuzzy c-means algorithm is a two-step iterative process that is defined as follows: First, set the input variables c (number of clusters equal to number of political parties), m (level of fuzziness), and ϵ (termination criterion, normally $\epsilon \in [0, 1]$). Second, set an iteration number $k=0$. Third, generate a matrix of cluster centers $\vec{Y}^{(k)}$ (refer to equation (5)) taking the mean average of answers from all candidate in the same political party (P_i). Then, given the initial matrix $\vec{Y}^{(k)}$, compute the fuzzy partition matrix $\vec{U}^{(k)}$ (refer to equation (4)).

Finally, using a repeat-until loop, update $\vec{Y}^{(k+1)}$ using $\vec{U}^{(k)}$ and then update $\vec{U}^{(k+1)}$ using $\vec{Y}^{(k+1)}$. Repeat this process until the termination criterion is reached ($|\vec{U}^{(k+1)} - \vec{U}^{(k)}| \leq \epsilon$).

The termination criterion could also be a predefined number of iterations. The modified fuzzy c-means is presented in Algorithm 2.

The outputs of the modified fuzzy c-means algorithm are: a fuzzy partition matrix $\vec{U}^{(k+1)}$ that contains the membership degree of voters (users) and candidates with respect to each cluster, and a matrix of cluster centers $\vec{Y}^{(k+1)}$. More details about fuzzy clustering analysis and the graphical interface implemented are presented in section 5.2.

The recommendation engine used for the experiments presented in this paper has the following input variables: number of clusters (number of political parties), number of closest candidates to be displayed, type of recommendation to be displayed, and voter (user) responses.

⁶Roles and Definition of Political Parties: <http://aceproject.org/ace-en/topics/pc/pca/pca01/pca01a>

Algorithm 2 FCM Modified

Input: c, m, ϵ
Output: $\vec{U}^{(k+1)}, \vec{Y}^{(k+1)}$

- 1: Set iteration number: $k \leftarrow 0$
- 2: **for** $i = 1$ to c **do**
- 3: $y_i \leftarrow$ mean average of answers from the i -th Political Party (P_i)
- 4: **end for**
- 5: Compute $\vec{U}^{(k)} \leftarrow \vec{Y}^{(k)}$
- 6: **repeat**
- 7: Update $\vec{Y}^{(k+1)} \leftarrow \vec{U}^{(k)}$
- 8: Update $\vec{U}^{(k+1)} \leftarrow \vec{Y}^{(k+1)}$
- 9: **until** $|\vec{U}^{(k+1)} - \vec{U}^{(k)}| \leq \epsilon$
- 10: **return** $\vec{U}^{(k+1)}, \vec{Y}^{(k+1)}$

Additionally, it uses the typical parameters of a FCM, $m = 2$, and $\epsilon = 1 * 10^{-4}$ (refer to Algorithm 1), and the Sammon mapping algorithm using PCA as initialization method, total iterations = 500 and relative tolerance = $1 * 10^{-9}$.

4.6 Top-N Recommendations

The top- N candidates similar to voter (user) v are generated by using the bi-dimensional profiles. The distances of all candidates, with respect to voter (user) v , are computed, and the N closest candidates are displayed. The similarity percentage ($S_{vc_i}(\%)$) of a voter (user) v and the i -th candidate (c_i) is computed using the most distant candidate or citizen (d_{max}) as a reference. The computation of similarity percentage is shown in (8):

$$S_{vc_i}(\%) = 100 - \left(\frac{100 * d_{vc_i}}{d_{max}} \right) \quad (8)$$

where d_{vc_i} is the distance between voter (user) v and the i -th candidate/citizen.

The outputs are the N closest candidates and their similarity percentage with respect to voter (user) v . More details about the top- N recommendations and the graphical interface implemented are presented in section 5.3.

4.7 Building Political Communities

The recommendation engine presented in this paper could also be used during eDiscussion and ePosting with the creation of so-called political communities, allowing citizens to interact through specific media, potentially crossing geographical and political boundaries in order to pursue mutual interests or goals. The use of the user-friendly bi-dimensional interfaces, could help voters (users) to establish which citizens are the most similar according to their preferences and tendencies (profiles).

To create political communities, the recommendation engine uses the datasets of citizens together with the dataset of candidates. In the first step, the recommendation engine transforms the high-dimensional profiles into a bi-dimensional space. Secondly, in order to compute the fuzzy clusters, only the bi-dimensional profiles of candidates and the voter (user) looking for the recommendation are used. Once the fuzzy clusters are computed, the datasets are merged (voter, citizens and candidates) and displayed in a bi-dimensional map. More details about the creation of political communities and the graphical interface implemented are presented

in section 5.4.

5. SmartParticipation Prototype

The SmartParticipation project uses a fuzzy recommender system prototype (FRSP) developed to display the results of a recommendation. Three output options were developed: fuzzy cluster analysis, top- N recommendations and community building tools, as shown in Fig. 4. The FRSP uses a dataset provided by the smartvote project, described in more detail in the next section, which corresponds to the Swiss national elections in 2007.

The FRSP is presented in Fig. 5 and has the following inputs:

- *Type of Recommendation:* This input is used to select the type of recommendation target.
- *Type of Relevance:* This input allows the user to select the type of relevance that will be used. This option was included for evaluation purposes, since the FRSP uses the dataset provided by smartvote and does not include relevance for candidates. Three options are included in this section: no relevance, which means that the relevance of candidates will be excluded from the dataset; same as voter option, which adds the same relevance of the subject citizen to all candidates on each question; and random relevance, which adds random relevance to each candidate for each question. In this project for simplicity, no relevance is used due to the fact that the other two options add noise to the dataset.
- *Voter Selection:* This input is used to select from the dataset of citizens who answered the smartvote questionnaire. For simplicity, only the citizens who answered all the questions have been taken into account.
- *Political Parties Selection:* This input allows the user to narrow down his/her recommendation for those political parties that are object of his/her interest.
- *Topics Selections:* This input allows the users to select only the topics that could be of interest.
- *Graphical Interface:* This input displays the selected type of recommendation.

In this section, the dataset used and the results obtained with the FRSP are presented and discussed.

5.1 Smartvote Dataset

The displayed results in this work correspond to the answers of candidates and voters (users) provided by smartvote project, an online VAA, for local, cantonal, and national elections in Switzerland.

The 2007 questionnaire consisted of more than 70 questions on the most important political issues (e.g., "Do you think that nuclear power plants should be shut down?"). The questions are divided into eleven groups related to specific topics, such as: welfare, family and health, education and sports, and migration and integration, among others. The possible answers to each question are "yes," "rather yes," "rather no," and "no." Citizens can decide to answer (or not) one or more questions. However, candidates do not have the possibility to opt out. They must answer all questions and

Figure 5: SmartParticipation Prototype

confirm their answers before they are saved in the smartvote database.

The dataset used in this paper contains the answers of 257 candidates to the two chambers of the Swiss parliament (i.e., the National Council and the Council of States), who responded to the 73 questions on the smartvote questionnaire. The results that are presented in this paper consider only the voters (users) who answered the complete questionnaire in order to provide a more accurate result.

The candidates, who are a part of the dataset, belong to the following political parties: the Central Democratic Union, Christian Democratic Party, Christian-Social Party, Evangelical Party, Federal Democratic Union, Green Party, Independent Citizens Movement, Labor Party, Left Alternative, Liberal Movement Ecology, Liberal Party, Opening Movement, Radical Democratic Party, Social Democratic Party, and Swiss Democrats.

The recommender system approach presented in this paper has three graphical interfaces: the fuzzy cluster analysis graphical interface (FCAGI), the top-N recommendations graphical interface (TNRGI), and the political communities graphical interface (PCGI), which are described in more detail in the following sections and presented in Figs. 6a, 6b, and 6c, respectively.

In the three graphical interfaces, the candidates and the voter (user) are displayed with different geometric figures. Each political party has a center represented by a geometric figure with the same shape and color of the figures representing the candidates belonging to that political party. The percentage of closeness of a voter (user) to the centers of each political party is presented as percentage.

In this paper, for simplicity, only three political parties are

taken from the smartvote dataset: The Federal Democratic Union (FDU), Radical Democratic Party (RDP), and Social Democratic Party (SDP). The selection of political parties was made taking into account that the political parties have a similar number of members.

5.2 Cluster Analysis

The FCAGI displays, in a bi-dimensional map, the locations of the voter (user) and the candidates (labeled by political parties), the clusters that are generated according to each political party, and the percentage of the closeness of the voter (user) to each cluster.

The fuzzy cluster analysis was completed by applying the bi-dimensional FP. Fuzzy clusters were generated using a modified version of the FCM algorithm (refer to section 4.5) with two main inputs: the number of clusters and a matrix of cluster centers. The number of clusters is equal to the number of political parties. The matrix of cluster centers was computed by taking the mean average of all answers from candidates in the same political party.

An example of FCAGI is presented in Fig. 6a. The displayed results indicate the formation of clusters with a clear concentration of candidates from the same political party.

Fig. 6a shows that the closest political party with respect to the voter (user) is the Federal Democratic Union (66%), followed by the Social Democratic Party (22%) and finally the Radical Democratic Party (12%).

In this experiment, some candidates apparently belong to different political parties. Fig. 6a clearly shows that two candidates who belong to the Social Democratic Party are closer to the cluster of the Federal Democratic Union, according to their profile (the answers given by candidates in

Figure 6: Recommendation output.

the smartvote system).

It is possible to explain this observation by realizing that the candidates who belong to a political party do not necessarily think in the same manner, which was an expected result.

5.3 Top-List of Neighbors

The TNRGI displays the location of a voter (user) and candidates (labeled by political parties), the clusters generated according to each political party (with a percentage of closeness of the voter (user) to each cluster), and the N closest candidates labeled with the percentage of proximity to the voter (user).

The top- N candidates who are similar to voter (user) v were generated using the bi-dimensional FP. The distances of all candidates, with respect to voter (user) v , were computed. The N closest candidates are displayed.

In order to generate the top- N candidates and similarity

percentages, the recommendation engine computes a vector of distances between the voter (user) and all candidates by using the normalized bi-dimensional profiles (refer to section 4.6)

Fig. 6b shows the results of the TNRGI with the same dataset used in previous section. It shows the formation of clusters by political party and the top-10 candidates close to the voter (user), together with the similarity percentages.

5.4 Range of Political Communities

In this section, an example of a potential political community is presented in Fig. 6c. In this figure, not only the candidates but also the citizens involved in the system are included in the bi-dimensional map. The citizens are represented by black squares, and for this experiment, the 20 closest citizens are represented by filled black squares.

For simplicity, the bi-dimensional map includes the citizens who answered all questions in the smartvote dataset.

After the generation of the bi-dimensional profiles, the clusters were generated with the voter (user) and the candidates' dataset. The reason for this, is that citizens are not defined to be part of any political party, and they cannot be used for the generation of cluster centers.

Finally, once the clusters were computed, the bi-dimensional datasets of citizens, candidates, and the voter (user) were merged and displayed. Fig. 6c shows the three political parties used in the previous experiments (refer to sections 5.2 and 5.3), the citizens who answered all questions from the smartvote questionnaire, and the voter (user) used in the experiments presented in sections 5.2 and 5.3.

Fig. 6c clearly shows that the clusters generated have a different shape, as compared to those in Fig. 6a and Fig. 6b. This can be explained by the nature of the Sammon mapping technique used by the recommendation engine which attempts to preserve inter-pattern distances (refer to section 4.4). For that reason, if more users are added to the dataset in the original space (high-dimensional), the resultant space (bi-dimensional) will differ.

6. Conclusions and Outlook

In this research, a recommender systems architecture for eGovernment, used in the SmartParticipation project, has been proposed. The Web-based recommendation engine can be used to visualize differentiated clusters of politicians as well as of citizens. It therefore supports collaboration, eElection processes for candidates, building processes for political communities that share common objectives, and civic participation.

The recommender system proposed can be used for eCollaboration, eDemocracy, and eCommunity. Based on a fuzzy clustering approach, it computes similarities between citizens and politicians in a multi-dimensional space. The Sammon mapping technique allows for a better understanding and evaluation of the relationships among citizens and/or politicians using a bi-dimensional graphical interface.

The creation of political communities and social networks among citizens allows for interaction and participation through social media, potentially crossing geographical and political boundaries. Contacting people with similar political profiles, building exchange platforms, and stimulating participation will enrich the information and knowledge-based society in the future.

The recommender system approach presented in this paper differs from collaborative filtering methods in that they are based on past experiences. It is also suitable in the one-and-only scenario, introduced by Guo and Lu[4], in which events such as election processes occur only once, and their participants (candidates and/or citizens) cannot be considered unique, since their presence at such events and their way of thinking can vary over time.

In future work, the SmartParticipation project could be used to evaluate whether candidates really act the way they claim they will. The FRS could display their location in the bi-dimensional map as candidates and show their moving positions during their political engagement as elected officials, allowing voters to easily understand politicians' behavior.

The recommender system approach presented in this paper, the fuzzy interface, and the graphical interfaces must be evaluated by citizens. In the case of the algorithms used, a comparison with different methods for dimensionality reduction and clustering algorithms will be performed.

7. References

- [1] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers Norwell, New York, 1981.
- [2] J. Fivaz and G. Felder. Added value of e-democracy tools in advanced democracies? the voting advice application smartvote in switzerland. In *Beyond eGovernment—Measuring Performance: A Global Perspective*, pages 109–122, Washington, DC, 2009. Public Technology Institute and ITEMS International.
- [3] J. Fivaz and D. Schwarz. Nailing the pudding to the wall: E-democracy as catalyst for transparency and accountability. In *International Conference on Direct Democracy in Latin America*, Buenos Aires, Argentina, March 2007.
- [4] X. Guo and J. Lu. Intelligent e-government services with personalized recommendation techniques. In *International Journal of Intelligent Systems*, volume 22, pages 401–417, March 2007.
- [5] A. Ladner, J. Fivaz, and J. Pianzola. Impact of voting advice applications on voters' decision-making. In *Internet, Politics, Policy 2010: An Impact Assessment*, Oxford, UK, September 2010. Oxford Internet Institute (OII, University of Oxford).
- [6] A. Meier. *eDemocracy & eGovernment - Maturity Levels of a Democratic Knowledge Society*. Springer-Verlag, Berlin, 2009.
- [7] K. Pearson. Principal components analysis. *The London, Edinburgh and Dublin Philosophical Magazine and Journal*, 6:566, 1901.
- [8] J. W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, c-18(5):401–409, May 1969.
- [9] D. Schwarz, L. Schädel, and A. Ladner. Pre-election positions and voting behavior in parliament: Explaining positional congruence and changes among swiss mps. *Swiss Political Science Review*, 16(4):533–564, 2010.
- [10] L. Terán, A. Ladner, J. Fivaz, and S. Gerber. *Fuzzy Methods for Customer Relationship Management and Marketing - Applications and Classification*, chapter 6. IGI Global, 2011. (forthcoming).
- [11] L. Terán and A. Meier. A fuzzy recommender system for e-elections. In *Electronic Government and the Information Systems Perspective*, volume 6267 of *LNCS*, pages 67–76, Bilbao, September 2010. Springer.
- [12] E. Vozalis and K. G. Margaritis. Analysis of recommender systems' algorithms. In *Proceedings of the Sixth Hellenic-European Conference on Computer Mathematics and its Applications - HERCMA 2003*, Athens, Greece, September 2003.
- [13] S. Walgrave, P. van Aelst, and M. Nuytemans. Do the vote test?: The electoral effects of a popular vote advice application at the 2004 belgian elections. *International Journal of Political Science*, 43(1):50–70, April 2008.
- [14] R. R. Yager. Fuzzy logic methods in recommender systems. *Fuzzy Sets and Systems*, 136:133–149, June 2003.
- [15] L. A. Zadeh. Fuzzy sets. *Journal of Information and Control*, 8:338–353, June 1965.

Acknowledgment

The authors would like to thank the members of the smartvote project (smartvote.ch), the members of the Fuzzy Marketing Methods Research Center (www.FMsquare.org), and the Information System Research Group at the University of Fribourg (diuf.unifr.ch/main/is) for contributing valuable thoughts and comments.

Special thanks to Andreas Ladner, Diana Pacheco, and Jan Fivaz for their support and help using the smartvote database to test the prototype.