JIS Accepted for Publication
By the Journal of Information Science: http://jis.sagepub.co.uk

JIS

Article

Profile-based recommendation: a case study in a parliamentary context

Journal of Information Science
1–20

DOI: 10.1177/0165551510000000
jis.sagepub.com
SAGE

Luis M. de Campos
Departamento de Ciencias de la Computación e Inteligencia Artificial. ETSI Informática y de Telecomunicación. CITIC-UGR. Universidad de Granada, Spain.

Juan M. Fernández-Luna
Departamento de Ciencias de la Computación e Inteligencia Artificial. ETSI Informática y de Telecomunicación. CITIC-UGR. Universidad de Granada, Spain.

Juan F. Huete
Departamento de Ciencias de la Computación e Inteligencia Artificial. ETSI Informática y de Telecomunicación. CITIC-UGR. Universidad de Granada, Spain.

Abstract
In the context of e-government and more specifically that of parliament, this paper tackles the problem of finding Members of Parliament (MPs) according to their profiles which have been built from their speeches in plenary or committee sessions. The paper presents a common solution for two problems: firstly, a member of the public who is concerned about a certain issue might want to know who the best MP is for dealing with their problem (recommending task); and secondly, each new piece of textual information that reaches the house must be correctly allocated to the appropriate MP according to its content (filtering task). This paper explores both these ways of searching for relevant people conceptually by encapsulating them into a single problem: that of searching for the relevant MP's profile given an information need. Our research work proposes various profile construction methods (by selecting and weighting appropriate terms) and compares these using different retrieval models to evaluate their quality and suitability for different types of information needs in order to simulate real and common situations.

Keywords
Information filtering; information retrieval; user profiles; parliamentary documents; content-based recommender systems

1. Introduction

There has been an exponential growth in web-published content in recent years [1]. As a result, not only has the task of finding useful information become even more important, but it is also more difficult than ever before. It is therefore increasingly harder for search engines to return relevant material to the users because of the amount of information that now exists and also the way that searches are conducted. In terms of search behaviour, the original active role of specifying the information needs of users to find what they were searching for has been complemented with many other search alternatives [2] [3]. These include the approach whereby users play a more passive role and the responsibility for filtering the desired content and removing any irrelevant results is left to the search system.

This last alternative is set in the context of content-based recommender/filtering systems [4] [5] which are able to select and recommend things such as songs, films or books to users on account of their tastes, thus reducing information overload. The "objects" to be recommended could present an almost infinite variety of typologies but in this paper we shall focus on people. For example, in a scientific context, we might be interested in discovering which other researchers are working on content-based filtering (recommendation task), or given a new scientific paper on recommender systems included a few seconds previously in an application such as ResearchGate, the application would determine the most suitable set of users according to their research fields and send it to them (filtering task). The first

Corresponding author:
Luis M. de Campos, Departamento de Ciencias de la Computación e Inteligencia Artificial. ETSI Informática y de Telecomunicación. CITIC-UGR. Universidad de Granada, 18014, Granada, Spain.
Email: lci@decsai.ugr.es

case can also be considered as a type of expertise retrieval or expert finding problem [6] [7] [8]. In both cases (recommendation and filtering) the research areas or interest topics must be expressed in such a way that relevant people can be found using this information, and there should then be some kind of repository for representing user tastes or preferences. This structure is called the profile [9] and in its simplest format, it stores a set of words that try to describe the user interests and preferences.

Although all the contributions presented in this paper could be extended to the problem of finding people in general, the framework for this research is e-government [10] and more specifically that of parliament. We have focused on recommending Members of Parliament (MPs) and filtering information to MPs because the intrinsic interest of the topic (see below) and also because we have been collaborating with the Parliament of Andalusia (Spain) for several years, developing an Information Retrieval System (IRS) for accessing its official documents [11]. One of the main goals of e-government is to simplify access for members of the public to the plethora of information generated by public administrations and to increase citizen involvement and keep them more informed (Government to Citizen Service [12]). In view of the vast amounts of information generated and received by national and regional parliaments, efforts need to be made in order to make the information discussed by MPs available to the general public. Efficient tools are required so that this information can easily be accessed by any user interested. But who are these users? They might be citizens, journalists, lobbies, parliamentary staff or even the MPs themselves. While each may have different information needs and use different search methods as stated in [13] [14] [15], they all need to access the information generated.

In this research, the "objects" to be found are MPs. For example, there are various reasons why a member of the public, a member of staff or a journalist might want to know which MPs are connected with the current Syrian refugee crisis, for instance, and this situation is clearly seen as a recommendation problem: the member of the public might want to write a letter to raise awareness about the problem, the member of staff might want to compile statistics about the MPs who have intervened in a parliamentary plenary session on the subject, and the journalist might want to know which MPs are in favour in order to support them. On the other hand, MPs who are working on agricultural committees in a regional chamber might be interested in documents produced by the EU Parliament dealing with subjects of this nature but not in others relating to culture or education, for example. Since this scenario clearly corresponds to a filtering problem, the Parliament's Information Technology Department could develop a filtering system to find those MPs interested in reading the latest document about the EU Common Agricultural Policy. This last example also reflects the need to filter and organize large volumes of information received by parliamentary staff and MPs [13]. Another interesting example where our approach could be used is the We the People[1] initiative, where any American citizen older than 13 can create an online petition to the Obama's Administration. This request is supported by other citizens who sign it. When a minimum of signatures is reached, the White House distributes it to the appropriate policy officials within the Administration, who respond it. This process is carried out manually by the White House staff but our approach could help them to find the most relevant officials much faster and, maybe accurately.

One contribution of this paper is to cast these two problems (i.e. recommending MPs and filtering information to MPs) into a more general information retrieval-based problem [16]. In this field, an IRS is fed by a set of documents (collection), which are represented in some form and then stored in an index for fast lookup. Users then formulate their information needs by means of a query and the IRS retrieves those documents which are more relevant (similar) to that query. Extrapolating this general IR philosophy to our context, the documents feeding the search engine would be textual representations by MPs, expressed in the form of profiles and extracted from their speeches in Parliament. The queries could be expressed by means of a short sentence (like the ones that we usually formulate in a general-purpose search engine such as Google), a couple of medium-length paragraphs (like a press release) or an entire MP's intervention in a parliamentary session (long). With this abstraction, we can solve the problem of finding relevant MPs given a query submitted by a citizen about refugees or finding MPs to whom to send a parliamentary debate about a new education reform act.

The concept of profile plays an important role in this paper. Although details of the structure, construction and use will be given in the following sections, it is important to mention that these are built automatically by considering the concept of a legislative initiative, i.e. the literal transcription of a parliamentary discussion about a request presented by an MP or political group. An initiative comprises its general information (title, key terms, date, etc.) plus the sequence of transcriptions of the MPs' speeches (including interventions, responses and rejoinders). An MP's profile is built from all her interventions in the initiatives in which she has participated and will contain the most representative words and their importance from the subjects in which she is usually interested. We believe that the transcriptions of the MPs speeches is a public, compact and reliable source of information about their interests.

The goal of this research is to show how these content-based filtering and recommending problems could be tackled using information retrieval techniques based on the concept of profile in an attempt to answer the following research questions:

(1) In the context of recommending/filtering MPs, is it better to build and use a collection of profiles (as we want to do) or simply to search using the raw collection of initiatives? In this last case the recommended MPs should be those participating in the top ranked initiatives.
(2) Is the weighting scheme used to select terms crucial for building a quality profile?
(3) Is the profile composition (in terms of which parts of speech are included) relevant for achieving increased performance?
(4) Is the number of terms comprising the profile important for effective recommending?
(5) Is there any difference between retrieval models in terms of performance?
(6) Is recommending performance affected by the size of the query formulated to the system?

The main contribution of this paper is to show that standard information retrieval techniques (general purpose search engines, classical term weighting methods, content-based profiles, etc.) can be used to deal with the interesting application proposed in this case study, namely developing techniques for both filtering documents to MPs and recommending MPs to citizens, journalists, lobbies, etc. Another contribution is the development of an extensive experimental study validating our approach and revealing how these techniques can be configured (appropriate weighting schemes, size and composition of the profiles,...) in order to obtain the best possible results.

The paper is organized as follows: Section 2 offers a general overview of the topics relating to our work; Section 3 shows how we build the profiles; Section 4 presents the evaluation and results obtained; and finally, Section 5 concludes the paper with general outcomes and further lines of research.

2. Related work

Generically, a recommender system [17] [18] is a piece of software that suggests items to users according to their tastes. It basically attempts to find the degree of affinity between a product and the user-specific information that describes his or her interests (see [19] for a contextualization of recommendation in the information search process). In the literature, we can find two main groups: collaborative filtering and content-based recommenders. While the first group is based on the concept of groups of users with similar tastes based on rating patterns [20], the second relies on the textual descriptions of the items to be recommended and on the concept of profile of the users' interests [4] [5] [21]. In our case, depending on the intended use of the proposed system, we either recommend documents to MPs or MPs to members of the public with a content-based approach.

Content-based recommender systems are usually built using two different techniques [18] [22]: the first generates recommendations heuristically using information retrieval-based methods [16] [23] [24] [25] and the second uses machine learning methods (mainly supervised classification algorithms for learning user interest models), e.g. nearest neighbours [26], rule induction [27], decision trees [28], neural networks [29], genetic algorithms [30] or naive Bayes [31]. Information retrieval-based methods and machine learning-based methods roughly correspond to the division proposed in [4] between filtering systems that follow the statistical concept and filtering systems that follow the knowledge-based concept, respectively. In our case, we have no information about irrelevant training documents, only about relevant documents such as each MP's speeches (we think that it is not safe to consider the speeches of other MPs as irrelevant documents for the target MP). When a machine learning perspective is used, therefore, positive unlabelled learning[2] [32] should be used. In any case, the approach proposed in this paper falls within information retrieval-based methods, where only content information is used, unlike those environments where a variety of social information sources could be used to recommend [33].

In terms of the application, many content-based (or hybrid) recommender systems exist for a wide range of domains: web pages [34], news [35], music [25], movies [36], books [37], emails [27], scientific literature [24], TV programmes [38], restaurants [39], tourism activities [40], museums [41], package holidays and tours [42], and courses and learning materials [43], etc.

There are also various e-government applications from the Government-to-Citizen (G2C) perspective: in [44], the authors propose a multi-agent system that enables not only government agency managers to design new services tailored to the citizen's needs and desires but also to recommend the most interesting services supplied by public administration offices to citizens according to their profile. Another citizen service recommender system is based on a hybrid technique using both content-based and social media information [45]. A fuzzy system is proposed in [46] that recommends candidates in an e-election process to voters according to voter preferences. From the Government-to Business (G2B) perspective, there are also various applications [22] although these are mainly based on collaborative rather than content-based recommender systems. We are unaware, however, of any recommender system in a parliamentary context such as the one we propose.

From an abstract point of view one of our problems (the recommendation task) has a great similarity with the expertise retrieval problem [7]: (i) instead of filtering documents the problem is to retrieve a ranked list of people; (ii) the skills or knowledge associated to them has not been given explicitly, but instead they have to be discovered from a set of documents. The state-of-the-art expertise retrieval algorithms are also content-based algorithms that use language model techniques [47] to rank people, being therefore highly related to document search [7]. These algorithms can be classified into two main approaches: document and candidate models [8]. Document model retrieves the documents that best describe the topic (query), and then considers the candidates that are associated with these documents as possible experts [8] [48]. The second model constructs a textual profile of each candidate from the documents with which this expert is associated. This profile is represented by a multinomial probability distribution over the vocabulary of terms. Then, a retrieval model is used to match both expert's profile and queries in order to obtain the desired ranking. In the literature, it can be found that the document method performs significantly better than the candidate model when there are sufficiently many associated documents per candidate [49].

Nevertheless, finding political actors has several peculiarities which make this problem somewhat different from expertise retrieval, and our experimental results seems to support this conclusion. One reason is that political actors are usually interested in a wider range of topics, as for example those related with the different committees where an MP usually participates and also those topics related to her constituency. So, it is most likely that only some parts of her skills are relevant to a particular user interest, and therefore a lot of noise can be included in the output ranking. Another source of difficulty is that MPs interests are dynamics and may evolve rapidly over time: new topics emerge every day and some of them might become obsolete after a given law is adopted, for instance. Anyway, we are also unaware of any expertise retrieval application in a parliamentary environment.

In the context of personalization [3], where the information seeking process is adapted to the user's context, profiles have been used to represent the topics in which she is interested in (gathered in any possible way) and to retrieve documents which are closer to the user [50]. In this case, the basic use of profiles is a kind of query expansion, where the original query formulated by the user is complemented with some terms contained in the profile, with the aim of guiding the search results to the user's interests [51].

Concerning user profile construction, in our case we want to represent MP interests and preferences. Two important aspects of this process are the acquisition of user information and the representation of the user profile. User information may either be collected explicitly based on user interrogation or implicitly by recording user behaviour in some way [4]. In many cases (and this is also the case with MPs), since users are reluctant to either provide any personal information or complete questionnaires, an explicit approach is usually unfeasible ([52] is an example where profiles are built from the information contained in social networks and [53] based on user-assigned tags). In our case, we shall use an implicit approach since we have a source of public information about MPs' interests: the transcriptions of their speeches when debating each initiative during the parliamentary debates. By mining these relevant documents we will extract the information necessary to build the profiles.

Regarding user profile representation, the four main approaches are sets of weighted keywords, semantic networks, weighted concepts and association rules, according to [4]. In [21], the authors distinguish between keyword-based and semantic analysis (which includes ontologies and encyclopaedic knowledge sources). Weighted keywords is the simplest and most common user profile representation and the easiest to build: they may be automatically learned from documents which are relevant to the user. In our case, since the source of knowledge about MPs is precisely the sets of documents containing the transcriptions of their speeches, we will adopt this approach. Examples of other work which also uses sets of weighted keywords can be found in [54] [55] [56] [57].

With regard to the time perspective, profiles can be static whereby they maintain the same information over time or dynamic in that they may be modified as user preferences change [4]. Our proposed profiles are static in the sense that they are not incrementally updated in time. In order to incorporate new information, our profiles must be re-built from scratch as new parliamentary debates discussing new initiatives arise. This does not represent a practical problem since the profile construction process is extremely fast.

3. Profile construction

In [4], profile construction comprises three stages: the first step is to collect user information, i.e. to determine what information would be useful and how this should be extracted; the second step is to find an adequate representation of the profile; and the third and final step is the construction itself.

In terms of the information that we shall consider in this paper to build MP's profiles, this will be based on their participation (speeches) in the initiatives discussed in Parliament during committee and plenary sessions and extracted from the corresponding records of proceedings, which are public documents, manually transcribed by documentalists. We assume that "you are what you speak", so that the speeches of an MP reflect her political interests. Basically, each speech given by the same i-th MP would be collected and combined in a document d_i, which would serve as the source for their profile. The entire set of documents, $D = \{d_1,...,d_m\}$ (where m is the number of MPs), will be the input for the profile construction stage. The underlying assumption of this approach is that if we create a document containing the transcriptions of all the initiatives where an MP has participated, we will have some very reliable evidence about the political interests of that MP: if all of their interventions are connected with education, then the profile would comprise education-related terms; if she participates on various committees covering different subjects (e.g. agriculture, culture, health), we can assume that these are the subjects in which she is interested and broadly speaking her profile would comprise different groups of terms relating to such committee topics.

We can consider profile representation to be a bag of words containing the terms extracted from the MP's set of speeches weighted according to their importance in the profile without taking account of term dependence. More formally, the profile of the i-th MP, P_i, is $P_i = \{(t_{i1}, w_{i1}),..., (t_{in}, w_{in})\}$ (where n is the number of terms included in the profile).

Another important aspect in profile construction from a grammatical point of view is the type of term that could comprise the profile. One common choice is to include all of them, removing only stop words (those without meaning, i.e. prepositions, articles, etc.). Another alternative might be to consider the part of speech of each word and maybe to consider only nouns in an attempt to simulate a concept profile, or to include verbs in order to consider actions. The type of word comprising the profile according to the selected part of speech could affect performance of the process conducted with it.

As the reader may guess, the construction of each profile is simply a process by which weights are computed for each document d_i and then the best *n* terms according to any criterion that considers these weights are selected (filtered) to become profile members. In this paper, three weighting schemes are used, the first two of which are based on well-known measures [58] [59] [60]:

(1) Tf → The frequency of terms in the intervention collection, d_i. The selection for each P_i is carried out by considering the most frequent terms for the i-th MP in order to capture their most common words.
(2) Tfidf → By considering the raw frequency of each term in the collection account is also taken of the rarity of the term within the entire collection. The idf of a term t, i.e. its inverse document frequency, is then computed as $\log(|D|/df_t)$, where df_t is the number of documents where term t occurs.
(3) Difference (Diff) → Introduced in [61] for personalizing purposes, its underlying idea is to consider how many times a term t appears in and outside the document d_i which contains all the speeches of the i-th MP, $f^+(t,d_i)$ and $f^-(t,d_i)$, respectively. If the first amount is greater than the second, then the term is useful for representing the i-th MP. If this is not the case, it is discarded. The Difference measure (Diff) of a term t in relation to the document d_i (i.e. the i-th MP) is the normalized frequency of that term in the document minus the normalized frequency outside that i-th document (i.e. in the documents associated to the other MPs) and is expressed by the following formula:

$$Diff(t, d_i) = \frac{f^+(t,d_i)}{f^+(d_i)} - \frac{f^-(t,d_i)}{f^-(d_i)} \quad (1)$$

Normalization is carried out by considering the total number of terms in d_i, $f^+(d_i)$, and outside d_i, $f^-(d_i)$, respectively. Once the measure has been computed for a given term t of a document d_i, if the value is greater than 0, then the term is representative to a certain degree of d_i because it is more frequent in d_i than in any other document and therefore suitable for inclusion in the profile. Otherwise, when it is less than 0, it is not representative.

At this point, each MP has an associated profile comprising the n most representative terms as extracted from their corresponding d_i, depending on the weighting scheme used. The profile would be ready for use as desired.

However, we are not going to use the profiles in its current form, as lists of weighted terms. Given a query submitted to the system (e.g. a request of a citizen or a new document to be filtered), we have to compare this query with the different MP profiles to select those which are most similar. We want to do that by using the capabilities of any standard search engine. For that reason we want to transform any profile (a list of weighted terms) into a normal document (a bag of words which does not include the term weights computed in the selection stage) which can be indexed and searched for using any IRS. To do that the terms in the profile are therefore replicated various times according to their weights, thereby simulating the number of times a term appears in a document and enabling the corresponding search engine to build its own weighting scheme. For this purpose, we have considered two possibilities:

(1) R-Tf → Replication of a term Tf times, i.e. the number of occurrences in d_i. This approach keeps the original distribution of the term. In this case, the computed weights are only used to select the terms in the profile.
(2) R-Prop → Replication proportional to the original weights in P_i, i.e. the higher the weight for a given term t, the greater the number of times it occurs in the profile. This approach is implemented using a linear transformation applied to the original weight, which calculates term occurrence in proportion to term weight. The term with the highest weight will have the maximum replication value, n, while the one with the lowest weight will have the minimum replication value, 1. The formula is as follows:

$$replications = (trunc)\left((n-1)x\,\frac{weight - minWeight}{maxWeight - minWeight} + 1\right), \quad (2)$$

where *maxWeight* and *minWeight* are the maximum and minimum weight values.

4. Evaluating profiles

With the general aim of evaluating the quality of the different profiles for the purposes of recommending MPs to users or filtering pieces of information to them, in this section of the article we shall describe the experimental settings and the results obtained in addition to the ensuing conclusions.

4.1. Collection and experimental parameter description

In order to carry out the evaluation, we first need a data set. In our case, this is derived from the collection of Records of Parliamentary Proceedings from the Andalusian Parliament[3] in Spain. Each document comprises a set of initiatives discussed in the corresponding session. We have selected all the 5,258 initiatives belonging to the 8^{th} term of office of this regional chamber. They contain a total of 12,633 different interventions marked up in XML. This collection[4], whose size is 148MB, contains a total of 136,209 paragraphs, 19,429,148 words and 73,443 unique terms (excluding stop words and stemming performed). In our experiments, we have only considered 132 MPs or technical guests who have spoken in at least 10 different initiatives (with less than 10 initiatives, we consider that there is not enough information to build a reliable profile). From the set of each MP's speeches we have learnt the profiles to test in this experimentation.

In terms of the different parameters to be considered, these can be classified according to where in the evaluation they participate:

(1) Profile construction:
   (1.1) Number of profile terms: 50, 250, 500, 750 and 1000, ranging from relatively small to larger profiles in order to test whether size is a relevant factor. The tokenization, stop word removal and stemming steps were carried out using the SpanishAnalizer facility of the Lucene library[5]. We used a standard Spanish stop word list, adding to it some other words which are particularly frequent in the parliamentary discourse but useless for retrieval purposes (as for example gentleman, "señoría" in Spanish).
   (1.2) Part of speech of the term: only nouns (N), only nouns and verbs (NV) and all types of words (A). The basic goal is to determine whether a profile comprising "concepts" (N or NV) is better than another containing all words (except stop words, which are removed). The extraction of the parts of speech was carried out using the Apache OpenNLP package[6].
   (1.3) Term selection method: Tf (t), Tfidf (i) and Diff (d).
   (1.4) Importance of the term in the profile: R-Tf (F) and R-Prop (P).

(2) Profile retrieval:
   (2.1) Search engine: the open source Lucene library[5].
   (2.2) Retrieval models: the Lucene implementations of the BM25, Language Model (LM) and Vector Space Model (VECT).
   (2.3) Input - Document collection: the set of 132 MPs' profiles built according to the previous parameter values. We shall experiment with 90 different document collections, corresponding to the 90 ways of constructing the MP profiles.
   (2.4) Input - Queries:
      (2.4.1) Source: In order to simulate different types of information needs in terms of size as mentioned in the introduction to this article, we shall consider long, medium and short queries. The first type exemplifies the case of filtering complete initiatives to MPs so the query source will be the initiative (qI). In the second case (medium queries), which represents for example a press release consisting of several paragraphs, we consider an automatic summary of each speech (qS) in the initiative. This summarization process has been carried out using the "More Like This" (MLT) Lucene query, which creates a frequentist summary, i.e. a new query selecting the most representative terms from a source, in our case the initiative. Finally, to model a typically short query submitted by a citizen to find a relevant MP, we shall use the initiative title (qT), i.e. a short description of its content.
      (2.4.2) Treatment: The first natural option is to consider the full text of the initiative, summary or title as a single query (SGL). A second alternative, however, assuming that we are faced with the problem of not being able to identify the special features of MPs' speeches with the SGL approximation, is to use the initiative structure to split the initial query into several subqueries, each grouping the text associated to the speeches of each MP who participated in the initiative discussion (the compound or CMP approach). Let us suppose that each subquery represents an MP's point of view in the initiative since it focuses on their own intervention. The rationale for this proposal is that each subquery may possibly and more accurately identify the MPs who participated in the initiative. We therefore hope that the compound query is more effective than the single query in those cases where both can be applied (only in the query initiative and summary). At a later stage, once all the results of each subquery have been obtained, a ranking fusion procedure [62] must be run in order to compute a unique ranking of MP profiles. Before the different lists are combined, and in order to avoid problems arising from different subquery lengths, the relevance value computed for each profile is divided by the score of the top MP's profile for a particular query, resulting in normalized scores for all subqueries which are therefore easily combinable. With regard to the combination, in this research we have implemented two good methods for ranking fusion as originally proposed in [63]:
         (i) MAX: a unique ranking is obtained by assigning the maximum of the scores to each MP profile and re-sorting.
         (ii) MNZ: the final score of each MP profile is computed by totalling the scores in each ranking, but promoting those profiles that appear more frequently by multiplying the final sum by the number of lists where the MP's profile appears.
   (2.5) Output: Ranking of MP profiles sorted according to their relevance degree.

(3) Performance evaluation:
   (3.1) Methodology: The set of initiatives is randomly partitioned into a training set (80%) and a test set (20%). The first set is used to build MP profiles and the second for the purposes of evaluation. This process is repeated five times, averaging the values of each round. In order to anonymise evaluation, we have removed information relating to MPs in the speeches in which they have been involved from the test set. We have also considered as ground truth the fact that a test initiative will only be relevant to those MPs who participate in it, thereby creating a rather conservative assumption since it is quite reasonable to assume that an initiative will also be relevant to other MPs.
   (3.2) Baseline: We have considered a situation which could serve as the baseline for our proposals, where no proper profiles are considered but MPs are recommended regardless. In this case, the documents comprising the collection are initiatives from the proceeding records. The system will index these initiatives and given a query (initiative, summary or title, depending on the case) will match it to these initiatives and rank them. We shall then assume that the MPs participating in each one are the right MPs

we are looking for. The comparison with this baseline will give us an idea of whether the use of profiles for these purposes is appropriate [64].

(3.3) Evaluation measures: As we shall measure the ability to predict whether the profile representing an MP is relevant for a given query, the system will return a ranking of MP profiles that best match it. More specifically, the search engine will return a total of ten. We focus on these top ten retrieved results because on average there are 2.4 interventions per initiative, and we want to know how our approach is able to find them in the top positions. Regarding the metrics, these are commonly used in IR, depending on what we wish to measure: in order to count how many MPs out of all the relevant MPs appear in the top ranking positions, we compute recall at top-10 MPs (rec@10); in order to measure the ranking quality, we use the Normalized Discounted Cumulative Gain [65] (NDCG@10); finally, as the number of relevant MPs varies greatly with the initiative, we shall also compute the Mean Average Precision (MAP@n_j) and the precision over the top n_j MPs (R-Precision), where n_j is the number of MPs participating in the j^{th} initiative.

In terms of presentation, the results will be organized according to query source. We shall therefore show the results related to the initiative as the query source (qI), the summary (qS) and finally the title (qT). Within each analysis we shall focus on finding the best parameter configuration and determining the best query treatment approach option (SGL vs CMP, whenever possible).

4.2. Querying with initiatives

We shall begin the study by considering the entire initiative (qI) as the query source. In this case, as previously explained, we have two alternatives: either to submit it as a single query (SGL) or a compound query (CMP).

4.2.1. Initiative as a single query

First of all, and taking into account the number of evaluation measures, we shall determine the degree of correlation between the four metrics used in the evaluation. Figure 1 shows the values obtained by each metric (y-axis) for each configuration (x-axis) for the BM25 retrieval model. The pattern is the same for the two other retrieval models, LM and VECT, so that we do not show them.

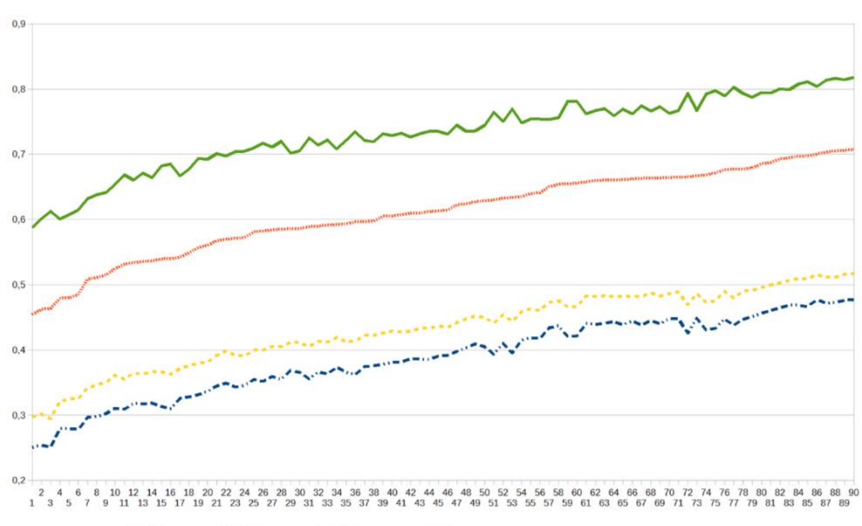

Figure 1. Correlations between evaluation metrics for the BM25 retrieval model.

A high correlation will have two main advantages for the development of this article: firstly, we will be able to reduce the number of results shown in this section by focusing on a single measure, thereby making this paper easier to read and understand; and secondly, the conclusions of this research will be strongly supported since the results will be consistent between the different metrics. In order to study the correlation, we computed the correlation coefficient between the metrics for the 90 parameter configurations for the three retrieval models. We obtained a minimum value of 0.96 and mostly a correlation of 0.99. This means that although each metric focuses on a single aspect, all four reflect the same tendency. As Figure 1 shows, the pattern is the almost the same for the four metrics. For reasons of space, we do not show the same graphics for CMP queries, summaries and titles but the tendency is exactly the same, which is why we shall use NDCG@10 as the main evaluation measure for the rest of the paper.

The second step would be to study the relative performance across retrieval models. In Table 1, we show the NDCG@10 values for the 90 possible configurations obtained with the BM25 model.

Table 1. NDCG@10 results for BM25 model. Bold values represent the highest values from each column. Parameters meaning: t: Tf; i: Tfidf; d: Diff; F: R-Tf ; P: R-Prop; N: only nouns; NV: only nouns and verbs; A: all types of words.

		50		250		500		750		1000	
		F	P	F	P	F	P	F	P	F	P
t	N	0.4543	0.4620	0.5356	0.5310	0.5675	0.5570	0.5897	0.5696	0.5930	0.5720
	NV	0.4800	0.4850	0.5810	0.5836	0.6129	0.6072	0.6325	0.6225	0.6409	0.6283
	A	0.5395	0.5396	0.6296	0.6336	0.6557	0.6543	0.6680	0.6652	0.6772	0.6715
i	N	0.5151	0.5249	0.5973	0.5966	0.6055	0.6050	0.6120	0.5912	0.6095	0.5710
	NV	0.5422	0.5489	0.6349	0.6396	0.6572	0.6604	0.6630	0.6608	0.6646	0.6595
	A	**0.5887**	**0.5965**	**0.6761**	**0.6772**	**0.6969**	**0.6971**	**0.7033**	**0.7048**	**0.7058**	**0.7074**
d	N	0.5084	0.5110	0.5921	0.5860	0.6142	0.6095	0.5853	0.5601	0.4792	0.4639
	NV	0.5337	0.5364	0.6273	0.6236	0.6541	0.6498	0.6641	0.6605	0.6669	0.6620
	A	0.5822	0.5851	0.6634	0.6629	0.6852	0.6789	0.6926	0.6874	0.7002	0.6945

In this table, the data are grouped into five main columns according to profile size (50, 250, 500, 750 and 1000 terms), and for each size two columns represent the two methods to reflect term importance in the profiles: R-Tf (F) and R-prop (P). The rows are split into three parts, corresponding to the three term selection approaches considered (Tf (t), Tfidf (i), and Diff (d)). Finally, each single row is related to the profile composition in terms of parts of speech (N, NV and A).

In order to enable conclusions to be drawn, the previous table is also expressed graphically in Figure 2, where we show the NDCG@10 values on the Y axis and the different profile compositions on the X axis (the terms A.B.C represent the PoS, the selection method and the profile weighting scheme, respectively), plotted according to the number of terms that they contain.

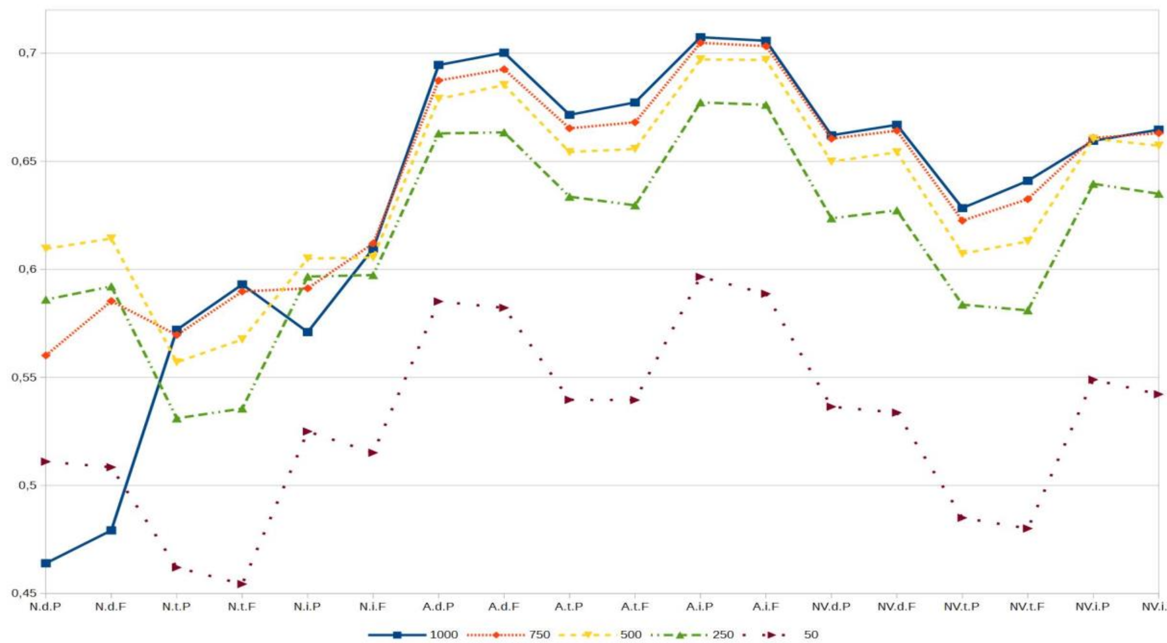

Figure 2. Performance (NDCG@10) for BM25 with different types of profiles plotted according to number of terms.

If we focus our attention on the best parameter configurations (i.e. those profiles that are best for recommending the correct MPs), we can observe from the plot in Figure 2 that there are two peaks in performance when the profiles comprise all parts of speech (A) and when Diff (d) and Tfidf (i) are used to select terms, respectively. This means that it is better not to consider any restriction on the selected terms, i.e. all parts of speech matter. The best choice appears to be Tfidf, although there are no significant differences with Diff. What is a fact is that these two approaches are better than the raw term frequency (t). If we now consider the two approaches for determining term weights in the profiles, i.e. proportional (P) and frequency-based (F), there is no clear superior method: there is no statistical significance between the results obtained with F and P (using a t-test).

We are not going to show the corresponding tables and figures for the two other retrieval models due to space considerations, but also because in general their results and tendencies are similar to those of BM25. The differences between BM25 and VECT are very low although BM25 is slightly better in high profile size values while VECT is generally better in medium and low values. LM shows a different pattern and worse results. The best values in these two first models are mostly located in the A.i.P (all PoS, Tfidf and R-Prop) configuration, while in LM, the best performance is found with profiles of the type A.d.F (all PoS, Diff and R-Tf). The model that obtains the absolute best result is BM25.

The last parameter in this study is the profile size. It is clearly apparent that the best results are obtained when the number of terms is 1000 (this is systematic across models). It is also clear that this increase is constant from 50 to 1000, a fact that leads us to ask whether this tendency is continuous until all the terms in the documents have been included in the profile. We therefore conducted a new experiment with the parameter configuration that performs better with larger profile sizes for the BM25 model. The graph in Figure 3 plots the results and these show how there was a constant and significant increase in performance until 1000 terms, after which it became more stable until 5000 terms and then decreased until all terms included in the profiles had been considered. We therefore believe that 1000 terms is a good number if we take into account the quality of the metric values and performance (the profiles would have a manageable size). This relatively high number of terms comprising the profile could lead to the conclusion that the use of profiles in personalization (where a low number of terms is recommended in order to avoid the query drift problem [52]) is a different task from the one at hand where we are looking for the best MP profile, and larger profiles are more useful when the query is also long. For the other retrieval models, the trends are basically the same.

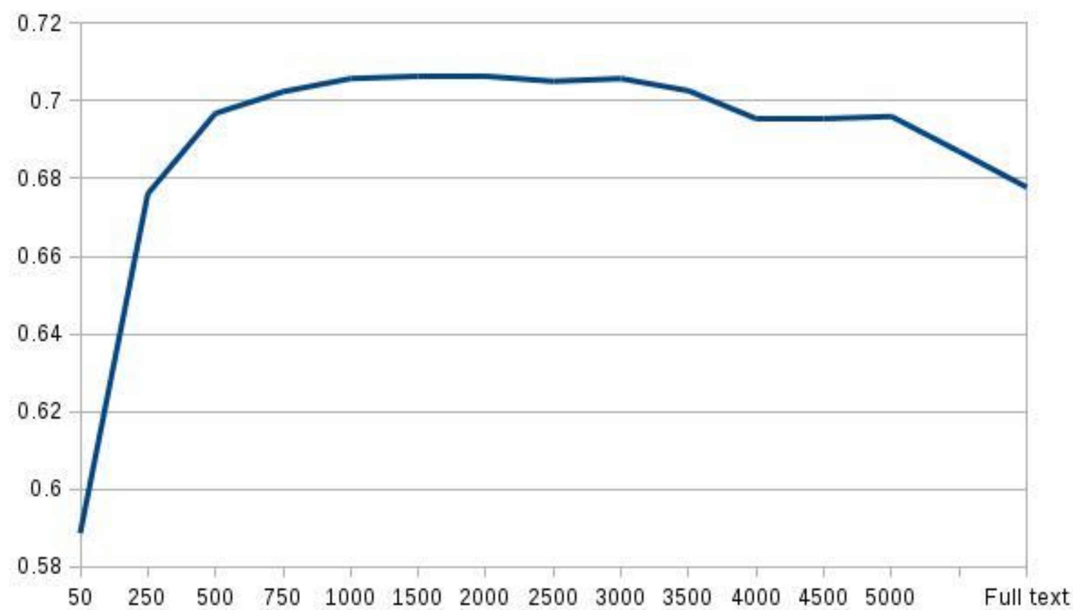

Figure 3. Representation of NDCG@10 values, using BM25, when varying the size of the profile.

Once we have presented the general performance of the profile recommendation, we wish to compare the results with those from the baseline method. In this case, we have considered a non-profile approach, where each document represents a complete initiative, as described in item 3.2 of Section 4.1. Table 2 shows the results for the best configurations of the profile-based approaches, comparing these with the baselines for each retrieval model. The improvement percentages are quite high, a fact which supports the use of profiles for the recommendation and filtering tasks presented in this paper. We have carried out a statistical analysis of these results, more precisely a t-test using the results on the five random partitions of the collection used in our experiments. The results confirm that all the detected differences between our models and the baseline are statistically significant, with p-values lesser than 0.01.

Table 2. Comparison between the baseline and profile-based approaches (best configurations) for the initiative as a single query (SGL).

Configuration Model	Size	Composition	Selection	Weight	NDCG@10	NDCG@10 baseline	% improvement
BM25	1000	All PoS	Tfidf	R-Prop	0.7074	0.5790	22.18
LM	1000	All PoS	Diff	R-Tf	0.6903	0.5750	20.05
VECT	1000	All PoS	Tfidf	R-Prop	0.7036	0.5277	33.33

In view of the discussions mentioned above and the fact that the performance pattern across retrieval models is practically the same in the other experiments (querying initiatives with compound queries, summaries and titles) as the one presented in this section, in the rest of the paper we shall only focus on BM25.

JIS Accepted for Publication
By the Journal of Information Science: http://jis.sagepub.co.uk

4.2.2. Initiative as a compound query

It should be remembered that in this CMP querying approach, the initiative (i.e. the SGL query) is split into various ones by grouping the text associated to the speeches of each MP participating in the initiative debate, with each comprising a subquery. A ranking is obtained for each query and these are fused by implementing the MAX and MNZ techniques which were presented previously.

Figure 4 shows a graphical representation of the NDCG@10 values for the two fusion methods used in CMP and for different profile sizes and the compositions of these for the BM25 model.

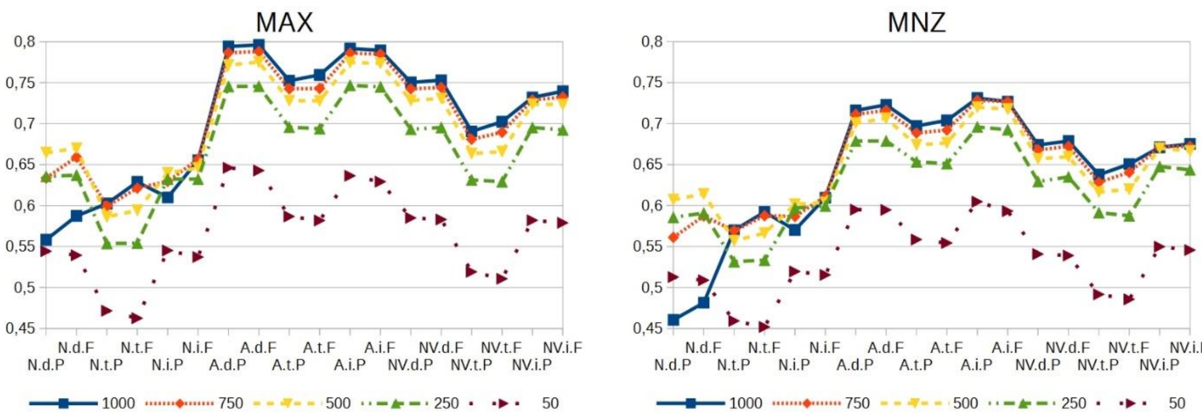

Figure 4. Performance for MAX and MNZ in CMP queries for BM25.

The first conclusion that can be drawn from the graphs is that this querying approach performs much better than the SGL one. In our opinion, submitting several subqueries that contain all of the MP's speeches is a far more precise reflection of their general line of discourse and is not diluted by other colleagues' speeches which may be about other subjects or even being longer. A second piece of evidence is that the way we aggregate the ranking matters since the MAX method outperforms MNZ in all configurations. It seems reasonable to think that preserving the best position for an MP's profile in a given ranking is much better than averaging its positions across all of them.

In terms of profile size, larger profiles perform best with 1000 again obtaining the best values. The differences between how different sizes perform are more noticeable. The patterns considering all the different parameters are similar to those shown in Section 4.2.1 and follow the same two peaks when all parts of speech are included in the profiles. It is preferable to select the terms with Diff and the choice of weighting scheme is immaterial as the differences between both are insignificant. The configurations that obtain the best values in this experimentation are presented in Table 3 together with the improvement percentage in relation to the baseline. Again, the corresponding t-tests indicate that the differences between the two CMP methods and both the baseline and the SGL method are statistically significant (p-values lesser than 0.01). The same happens with the difference between MAX and MNZ.

Table 3. Comparison between the baseline and profile-based approaches (best configurations) for the initiative as a compound query for BM25.

Querying - Fusion methods	Size	Composition	Selection	Weight	NDCG@10	NDCG@10 baseline	% improvement
CMP - MAX	1000	All PoS	Diff	R-Tf	0.7961	0.5790	37.50
CMP - MNZ	1000	All PoS	Tfidf	R-Prop	0.7310	0.5790	26.25

Finally, we present a graphical representation in Figure 5 of a comparison between the two querying approaches presented so far: SGL and CMP (for the profiles with 1000 terms). The best option is clearly a compound query using MAX as a fusion method. The SGL alternative is similar to MNZ in some cases, although worst in most.

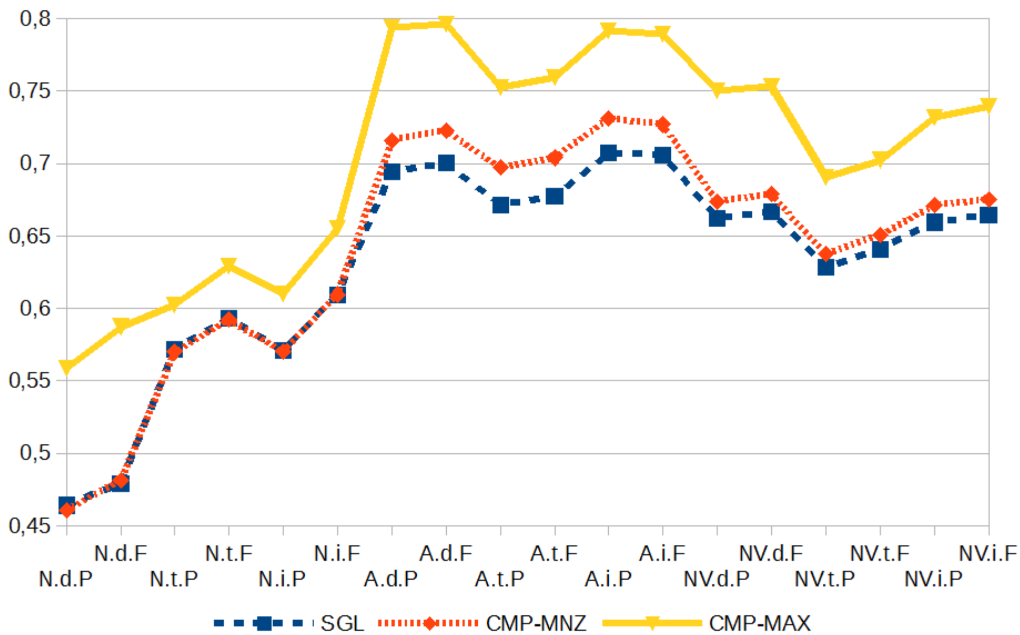

Figure 5. Performance for SGL and CMP queries with 1000 terms in the profile

4.3. Querying with summaries

In order to simulate medium-sized queries, as explained in Section 4.1, we have used Lucene's "more-like-this" (MLT) frequentist summarization (qS). This technique computes the term frequency of each term from an initiative, combining it with an idf-like measure and selects the top terms with the highest weights in order to represent the document. By using summaries, we can generate SGL and CMP queries and so in this section, we shall compare how they perform and select the top 25 terms for MLT. Figure 6 presents the graphical representation of the evaluation measure for SGL, CMP (MAX and MNZ) queries.

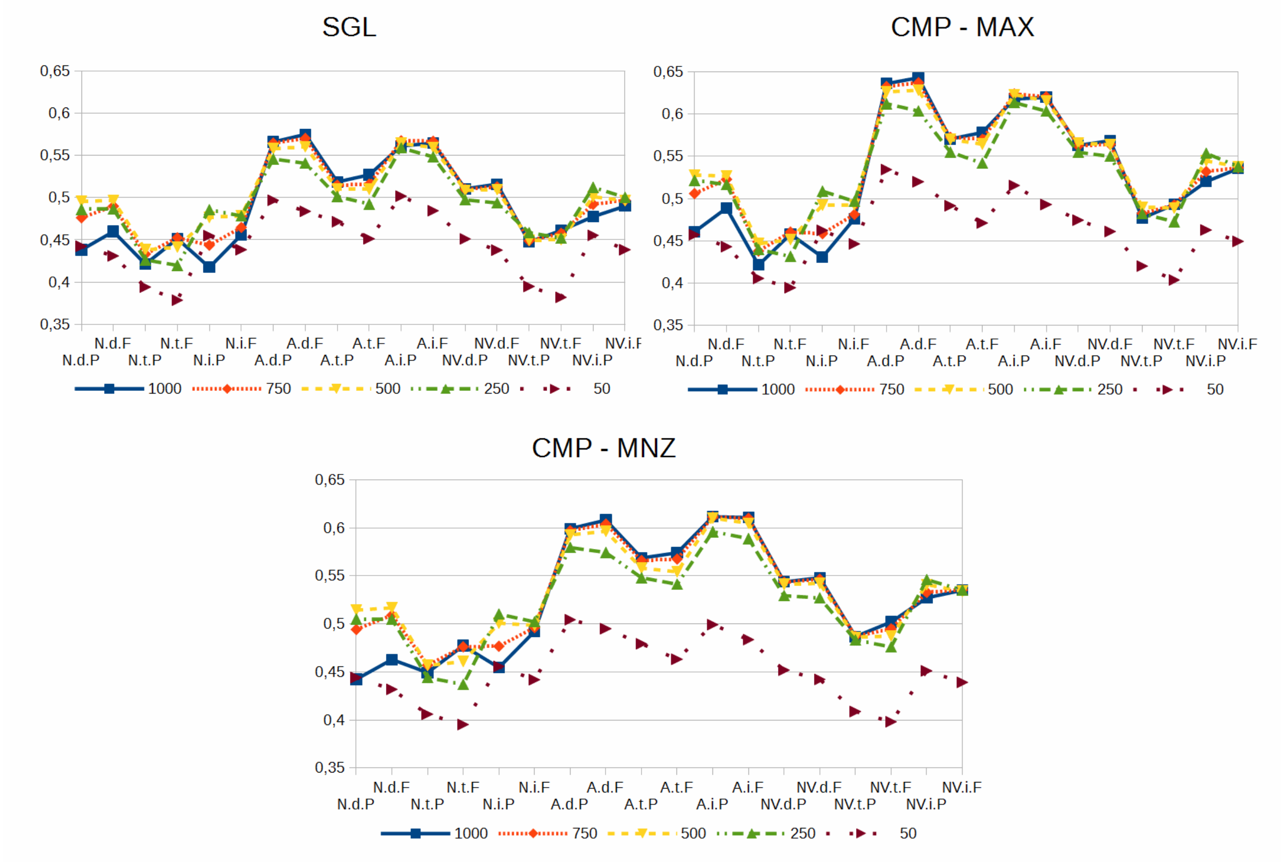

Figure 6. Performance for qS queries: SGL vs CMP (MAX and MNZ).

It is clearly apparent that the CMP approach is better than the SGL one, as occurred with qI, with MAX being the best ranking fusion method. It is also evident that there has been a drop in performance in relation to considering the initiative as the query source.

In terms of profile composition, the profiles containing all PoS are still the best choice as these obtain the highest values in the three evaluated methods. The combination A.d.F (all PoS, Diff and R-Tf) offers the best results for SGL and MAX in larger profile sizes while A.i.P (all PoS, Tfidf and R-Prop) is better in smaller ones. For MNZ, the best combination for all sizes is constantly A.i.P. For all PoS, 1000 terms is generally the size with the highest values although the difference in relation to a size of 750 is very low. Figure 7 compares these three evaluated alternatives with 1000 terms and shows how MAX performs better than MNZ and both perform better than SGL.

Journal of Information Science, 2016, pp. 1-20 © The Author(s), DOI: 10.1177/0165551510000000

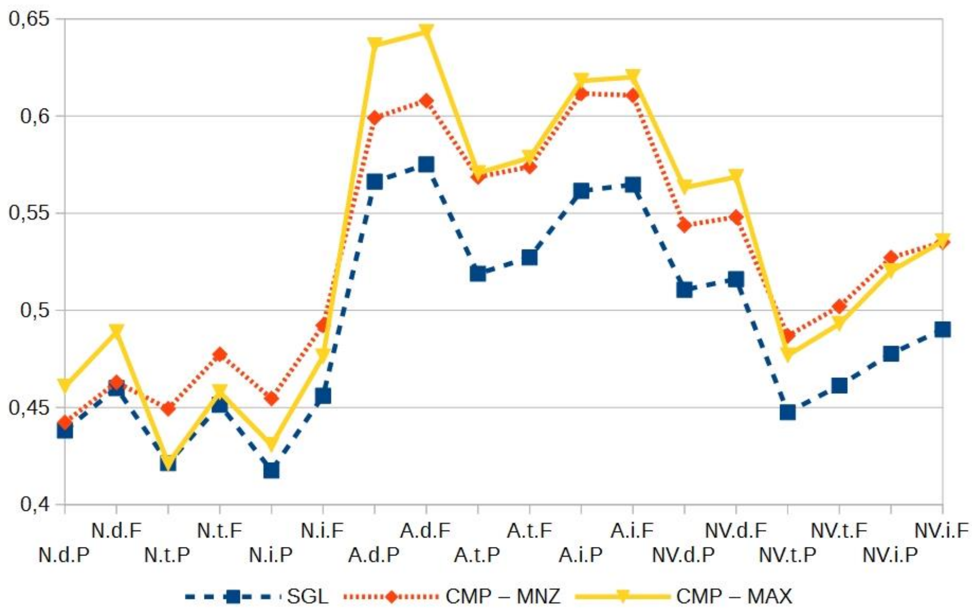

Figure 7. Performance for qS queries with 1000 terms in the profile.

Table 4 contains the parameter combination that offers the best NDCG@10 values and a comparison with their corresponding baseline. This comparison is still positive for CMP, although smaller than for qI. All the differences with the baseline are statistically significant (with the p-values lesser than 0.01 in the cases of SGL and MAX, and lesser than 0.05 in the case on MNZ). Comparing CMP using qI and qS, we should mention that the trends are similar, configuring a consistent query mechanism.

Table 4. Comparison between the baseline and profile-based approaches (best configurations) for the summaries as single and compound queries for BM25.

Querying - Fusion methods	Size	Composition	Selection	Weight	NDCG@10	NDCG@10 baseline	% improvement
SGL	1000	All PoS	Diff	R-Prop	0.5751	0.5967	-3.62
CMP - MAX	1000	All PoS	Diff	R-Tf	0.6433	0.5967	7.80
CMP - MNZ	1000	All PoS	Tfidf	R-Prop	0.6116	0.5967	2.50

4.4. Querying with titles

Finally, and to end this experimentation section, we have used the initiative title (qT), which has a relatively low number of terms and could be assimilated to those queries that a user usually submits to a search engine in order to simulate small-sized queries as explained in Section 4.1. In this case, there is no possibility of using compound queries (CMP) as the title is an initiative attribute which is common for all interventions.

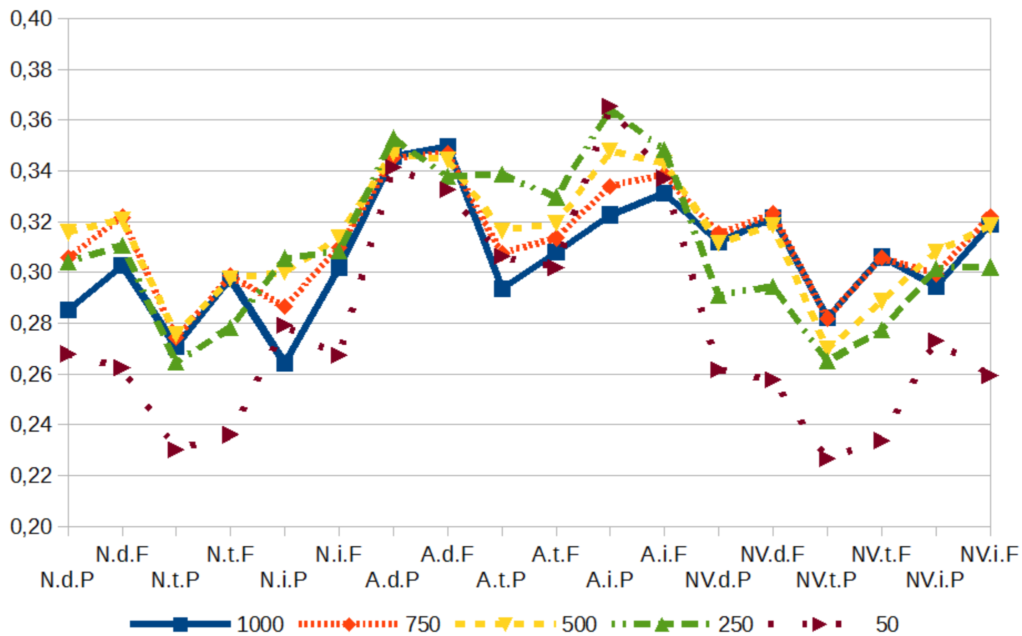

Figure 8. Performance for qT queries.

The results depicted in the graph in Figure 8 enable us to draw two clear conclusions: firstly, performance in comparison with the qI querying approach is extremely low; and secondly, it is a matter of fact that the profiles containing all the PoS are better, and for these, Diff offers greater NDCG@10 values than Tfidf in larger profiles and conversely in the smaller ones. We may also conclude that the profile size is relevant, breaking the trend of 1000 terms as the best profile size in qI and qS. The best absolute value can be found in profiles with 50 terms. It appears that we need to use much smaller profiles with fewer terms when reducing the size of the query as larger profiles could introduce a lot more noise, thereby impairing recommendation performance. It is also interesting to note that there is a break in the tendency found in qI and qS whereby the larger the profile, the better the performance, where it is not easy to find a clear pattern.

Table 5 shows the best configuration for qS and the comparison with its corresponding baseline, where we observe low performance (statistically significant difference, with p-value lesser than 0.01).

Table 5. Comparison between the baseline and profile-based approach (best configuration) for initiative titles as a single query for BM25.

Query source	Size	Composition	Selection	Weight	NDCG@10	NDCG@10 baseline	% improvement
qI	50	All PoS	Tfidf	R-Prop	0.3654	0.3871	-5.60

5. Concluding remarks and future work

In this paper, we have presented a way of using profiles either to recommend MPs to users or to filter documents to MPs. These are built using the initiatives in which each MP has participated. Our approach is supported by the fact that profiles are used as documents from a collection which is indexed by a search engine. Given a query, the retrieval system returns a ranking of MP profiles. This work has served to explore various alternatives for creating the profiles: composition in terms of parts of speech, different term selection methods, term weighting schemes, and finally the size of these structures. Additionally, we have simulated three ways of expressing information needs, which correspond to large, medium and short queries. These sizes could correspond to full documents or press releases, which could be filtered to the relevant MPs or user-formulated queries in order to find MPs.

A general but important conclusion of this work is that profiles are appropriate structures for supporting recommendation and filtering tasks in the parliamentary context where this paper is set. This contrast with the situation found in expert retrieval, where document models not using profiles (as the baseline approach that we considered in this paper) perform better than candidate models which do use profiles. In our parliamentary context the opposite occurs. Another difference, although is based on results not reported in the paper, is that those successful voting mechanisms in document-based model for expert search (which aggregate information for the top ranked documents in order to select the best experts) do not work well in our context, where it is better to consider only the maximum of the scores of the documents.

More specific conclusions from the analysis of the empirical results presented in Section 4, linked as answers to the research questions proposed in Section 1, are as follows: First, we have shown how representing the MPs' interests by means of profiles, in most of the situations improves the performance in the tasks at hand, in opposition to use a raw collection of initiatives, which has served as baseline in this paper (Research Question no. 1, RQ#1). Regarding the weighing schemes, we consider that the correct selection of an appropriate scheme is important for building quality profiles. Moreover, it is preferable to select the terms that will comprise the profile by taking into account measures that simultaneously contemplate information from the document itself and the collection, such as the case of Diff and Tfidf. With regard to the weights used once the terms are selected to be part of the profile and to measure their importance, both R-Prop and R-Tf are valid alternatives since there are no significant differences between them to make one preferable to the other (RQ#2). Focusing on the type of term contained in the profiles (RQ#3), from a grammatical point of view, it is better to create profiles with all parts of speech (once stop words have been removed) as they best represent the concepts expressed in the speeches. Answering RQ#4 about the number of terms comprising the profiles, we have to say that, in general, relatively large profiles (about 1000 terms) perform well for large and medium queries. This is interesting because when they are used in personalization tasks, the recommendation is to keep them small. For small queries it is better to use a lot fewer terms. This is also a logic conclusion if we consider the fact that if the profile is large and the query small, the level of noise is very high and the retrieval performance is negatively affected. BM25 and VECT models are very similar in terms of retrieval performance and configuration patterns. LM behaves differently and performs slightly worse (RQ#5). Finally for RQ#6, we have to conclude that larger queries obtain better results than medium-sized ones, which in turn are better than smaller ones. In particular, the results with short queries did not improve the baseline approach. This means that if we wish to find MPs it is better to supply as much as information as possible about our information need to obtain the most suitable MP. In the case of short queries, perhaps they should be enlarged using some type of query expansion or other query modification technique.

As explained in the introduction to this paper, although the context of this paper is a parliamentary setting, the findings revealed in this study could be applied to other situations where the main objective would be to recommend relevant people.

Our future main lines of research will continue this work by studying the creation and use of sub-profiles in the context of the MP recommendation/filtering problem. In this parliamentary environment, MPs could deal with several matters and their intervention speeches might cover different subjects which would be included in the same profile. With the sub-profile approach we could create, for example, two different profiles: one with education-related terms and the other with health-related terms. Our hypothesis is that recommendation could be more effective. Another research line is guided by the fact that user interests change with time. This means that we could work on the creation of long and short-term profiles in order to represent historical interests and more specific ones, respectively. The profiles that we have considered in this work are keyword-based; we also want to explore other profile representations, especially those based on concepts. Finally, in this paper we have tackled the MP recommendation/filtering problem from an IR perspective; we are also interested in exploring the problem from the alternative machine learning approach (particularly using positive unlabelled learning).

Notes

1. http://petitions.whitehouse.org

2. A type of binary classification problem where only a set of positive examples is available, together with a larger set of unlabelled examples, and there is no set of negative examples.
3. http://www.parlamentodeandalucia.es
4. http://irutai2.ugr.es/ColeccionPA/legislatura8.tgz
5. http://lucene.apache.org
6. https://opennlp.apache.org

Funding

This research has been supported by the Spanish "Ministerio de Economía y Competitividad" under project TIN2013-42741-P and the European Regional Development Fund.

References

[1] Gantz JF. The diverse and exploding digital universe. IDC white paper 2008; 2:1-16.
[2] Case DO. Looking for information: A survey of research on information seeking, needs, and behavior. Academic Press, 2002.
[3] Ghorab MR, Zhou D, O'Connor A and Wade V. Personalised information retrieval: survey and classification. User Modeling and User Adaptation Interaction 2013; 23:381–443.
[4] Hanani U, Shapira B and Shoval P. Information filtering: Overview of issues, research and systems. User Modelling and User-Adapted Interaction 2001; 11:203-259.
[5] Pazzani M and Billsus D. Content-based recommendation systems. In: The Adaptive Web, Lecture Notes in Computer Science, Vol. 4321, 2007, pp. 325-341.
[6] Hofmann K, Balog K, Bogers T and de Rijke M. Contextual factors for finding similar experts. Journal of the American Society for Information Science and Technology 2010; 61(5):994-1014.
[7] Balog K, Fang Y, de Rijke M, Serdyukov P and Si L. Expertise Retrieval. Foundations and Trends in Information Retrieval 2012; 6(2-3):127-256.
[8] Balog K, Azzopardi L and de Rijke M. Formal models for expert finding in enterprise corpora. In: Proceedings 29th ACM SIGIR Conference on Research and Development in Information Retrieval. 2006, pp. 43-50.
[9] Gauch S, Speretta M, Chandramouli A and Micarelli A. User profiles for personalized information access. In: The Adaptive Web, Lecture Notes in Computer Science, Vol. 4321, 2007, pp. 54-89.
[10] Marchinonini G, Samet H and Brandt L. Digital Government. Communications of the ACM 2003; 46:25-27.
[11] de Campos LM, Fernández-Luna JM, Huete JF, Martín-Dancausa C, Tagua-Jiménez A and Tur-Vigil C. An integrated system for managing the Andalusian parliament's digital library. Program-Electronic Library and Information Systems 2009; 43(2):156-174.
[12] Palvia S and Sharma S. E-Government and E-Governance: Definitions/Domain Framework and Status around the World. In: Foundations of E-Government. 5Th International Conference of E-Governance, 2007. http://www.iceg.net/2007/books/1/1_369.pdf (2007, accessed November 2015).
[13] Svarre T and Lykke M. Professional e-government seeking behaviour. In Proceedings of the American Society for Information Science, 2014; 50: Article No. 23.
[14] Terra A. European information users. Topics from a comparative analysis within the European Union. New Library World 2010; 111:189-202.
[15] Marcella R, Baxter G and Davies S. The information needs and information-seeking behaviour of the users of the European Parliament Documentation Centre. Journal of Documentation 2007; 63:920-934.
[16] Baeza-Yates R and Ribeiro-Neto B. Modern Information Retrieval. Addison-Wesley, 2011.
[17] Resnick P and Varian H. Recommender systems. Communication of the ACM 1997; 40(3):56-58.
[18] Bobadilla J, Hernando A, Fernando O and Gutiérrez A. Recommender systems survey. Knowledge Based Systems 2013; 46:109-132.
[19] Furner J. On recommending. Journal of the American Society for Information Science and Technology 2002; 53(9):747-763.
[20] Cacheda F, Carneiro V, Fernández D and Formoso V. Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. ACM Transactions on the Web 2011; 5(2):1-33.
[21] Lops P, de Gemmis M and Semerano G. Content-based recommender systems: State of the art and trends. In: Recommender Systems Handbook, Springer, 2011, pp. 73-105.
[22] Lu J, Wu D, Mao M, Wang W and Zhang G. Recommender system application developments: a survey. Decision Support Systems 2015; 74:12-32.
[23] Belkin NJ and Croft WB. Information filtering and information retrieval: Two sides of the same coin? Communications of the ACM 1992; 35:29-38.
[24] Foltz P and Dumais S. Personalized information delivery: an analysis of information filtering methods. Communications of the ACM 1992; 35:51-60.
[25] Loeb S. Architecting personal delivery of multimedia information. Communications of the ACM 1992; 35:39-48.

[26] Billsus D, Pazzani M and Chen J. A learning agent for wireless news access. In: Proceedings of the International Conference on Intelligent User Interfaces, 2002, pp. 33-36.
[27] Cohen W. Learning rules that classify e-mail. In: Papers from the AAAI Spring Symposium on Machine Learning in Information Access, 1996, pp. 18-25.
[28] Kim J, Lee B, Shaw M, Chang H and Nelson W. Application of decision-tree induction techniques to personalized advertisements on internet storefronts. International Journal of Electronic Commerce 2001; 5:45-62.
[29] Jennings A and Higuchi H. A user model neural network for a personal news service. User Modelling and User-Adapted Interaction 1993; 3:1-25.
[30] Tjoa AM, Hofferer M, Ehrentraut G and Untersmeyer P. Applying evolutionary algorithms to the problem of information filtering. In: Proceedings of the 8th International Workshop on Database and Expert Systems Applications, 1997, pp. 450-458.
[31] Pazzani M and Billsus D. Learning and revising user profiles: the identification of interesting web sites. Machine Learning 1997; 27:313-331.
[32] Li X and Liu B. Learning to classify texts using positive and unlabelled data. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence, 2003, pp. 587-592.
[33] Lai C, Liu D and Liu M. Recommendation based on personalized tendency for different aspects of influences in social media. Journal of Information Science 2015, 41(6):814-829.
[34] Asnicar FA and Tasso C. ifWeb: a prototype of user model-based intelligent agent for document filtering and navigation on the World wide Web In: Proceedings of the Workshop Adaptive Systems and User Modelling on the World Wide Web, 6th International Conference on User Modelling, 1997, pp. 3-11.
[35] Mobasher B, Cooley R and Srivastava J. Automatic personalization based on Web usage mining. Communications of the ACM 2000; 43:142-151.
[36] Salter J and Antonopoulos N. CinemaScreen recommender agent: combining collaborative and content-based filtering. IEEE Intelligent Systems 2006; 21:35-41.
[37] Mooney RJ and Roy L. Content-based book recommending using learning for text categorization. In: Proceedings of the fifth ACM Conference on Digital Libraries, 2000, pp. 195-204.
[38] Smyth B and Cotter P. A personalized television listings service. Communications of the ACM 2000; 43:107-111.
[39] Hung-Wen T and Von-Wun S. A personalized restaurant recommender agent for mobile e-service. In: Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service, 2004, pp. 259-262.
[40] Moreno A, Valls A, Isern D, Marin L and Borras J. SigTur/E-destination: ontology-based personalized recommendation of tourism and leisure activities. Engineering Applications of Artificial Intelligence 2013; 26:633-651.
[41] Ruotsalo T, Haav K, Stoyanov A, Roche S, Fani E, Deliali R, Makela E, Kauppinen T and Hyvönen E. SMARTMUSEUM: a mobile recommender system for the web of data. Journal of Web Semantics 2013; 20:50-67.
[42] Schiaffino S and Amandi A. Building an expert travel agent as a software agent. Expert Systems with Applications 2009; 36:1291-1299.
[43] Lu J. A personalized e-learning material recommender system. In: Proceedings of the 2nd International conference on Information Technology and Applications, 2004, pp. 374-379.
[44] De Meo P, Quattrone G and Ursino D. A decision support system for designing new services tailored to citizen profiles in a complex and distributed e-government scenario. Data and Knowledge Engineering 2008; 67:161-184.
[45] Abdellatif A, Ben Amor N and Mellouli S. An intelligent framework for e-government personalized services. In: Proceedings of the 14th Annual International Conference on Digital Government Research, 2013, pp. 120-126.
[46] Terán L and Meier A. A fuzzy recommender system for eElections. In: Electronic Government and the Information Systems Perspective, 2010, LNCS, Springer, 6257, pp. 72-76.
[47] Zhai C and Lafferty J. A study of smoothing methods for language models applied to information retrieval. ACM Transactions on Information Systems 2004; 22(2):179–214.
[48] Macdonald C and Ounis I. Voting for candidates: adapting data fusion techniques for an expert search task. In: Proceedings of the 2006 ACM CIKM International Conference on Information and Knowledge Management, 2006, pp. 387–396.
[49] Balog K, Bogers T, Azzopardi, L, de Rijke M and van den Bosch A. Broad expertise retrieval in sparse data environments. In: SIGIR, 2007, pp. 551–558.
[50] Pasi G. Issues in personalizing information retrieval. IEEE Intelligent Informatics Bulletin 2010; 11(1): 3-7.
[51] de Campos LM, Fernández-Luna JM, Huete, JF and Vicente-López E. Using Personalization to Improve XML Retrieval. IEEE Transactions on Knowledge and Data Engineering 2014; 26 (5):1280-1292.
[52] Mazhari S, Fakhrhmad S and Sadeghbeygi H. A user-profile-based friendship recommendation solution in social networks. Journal of Information Science 2015; 41(3):284-295.
[53] Mao J, Lu K, Li G and Yi M. Profiling users with tag networks in diffusion-based personalized recommendation. Journal of Information Science 2015; DOI: 10.1177/0165551515603321.
[54] Ahn J, Brusilowsky P, Grady J, He D and Syn SY. Open user profiles for adaptive news systems: help or harm? In: Proceedings of the 16th International Conference of World Wide Web, 2007, pp. 11-20.
[55] Lieberman H. Letizia: an agent that assists web browsing. In: Proceedings of the 14th International Join Conference on Artificial Intelligence, 1995, pp. 924-929.

[56] Moukas A. Amalthaea: information discovering and filtering using a multiagent evolving ecosystem. Applied Artificial Intelligence 1997; 11:437-457.
[57] Pazzani M, Muramatsu J and Billsus D. Syskill & Webert: identifying interesting web sites. In: Proceedings of the 13th National Conference on Artificial Intelligence, 1996, pp. 54-61.
[58] Salton G and Buckley C. Term-weighting approaches in automatic text retrieval. Information Processing & Management 1988; 24(5):513-523.
[59] Sparck Jones K, Walker S and Robertson SE. A probabilistic model of information retrieval: development and comparative experiments. Information Processing & Management 2000; 36(6):779–808.
[60] Sparck Jones K, Walker S and Robertson SE, A probabilistic model of information retrieval: development and comparative experiments (part 2), Information Processing & Management 2000; 36(6):809–840.
[61] de Campos LM, Fernández-Luna JM, Huete JF and Vicente-López E. XML Search Personalization strategies using query expansion, reranking and a search engine modification. In: Proceedings of the ACM Symposium on Applied Computing, 2013, pp. 872-877.
[62] Wu S. Data fusion with estimated weights. In: Proceedings of the 2002 ACM CIKM International Conference on Information and Knowledge Management, 2002, pp. 648-651.
[63] Fox E and Shaw J. Combination of multiple searches. In: Proceedings of the Second Text REtrieval Conference (TREC-2), 1994, pp. 243-252.
[64] de Campos LM, Fernández-Luna, JM and Huete JF. A lazy approach for filtering parliamentary documents. In: Proceedings of the 4th International Conference on Electronic Government and the Information Systems Perspective, 2015, LNCS 9265, pp. 364-378.
[65] Jarvelin K and Kekalainen J. Cumulative Gain-based Evaluation of IR Techniques. ACM Transactions on Information Systems, 2002; 20: 422-446.