Vietnam J Comput Sci (2016) 3:47–56
DOI 10.1007/s40595-015-0054-9

CrossMark

REGULAR PAPER

$\mathcal{IRORS}$: intelligent recommendation of RSS feeds

Nedia Araibi[1] · Eya Ben Ahmed[2] · Wahiba Karaa Ben Abdessalem[1]

Received: 20 September 2015 / Accepted: 28 November 2015 / Published online: 14 January 2016

Abstract The abundance of information prohibits getting relevant results on online social researches. Thus, RSS feeds appear as monitoring tool of current events according to users preferences. However, the user is flooded by the amount of such RSS feeds. For that reason, any analysis of RSS feeds seems effortful and complex. In this paper, we aim to improve the effectiveness and swiftness of pertinent RSS feeds analysis through recommending suitable fragments of queries during the analysis process of events. Accordingly, we propose an innovative architecture of our new active RSS feeds warehouse. Additionally, we introduce a new recommender system to improve the querying expression of RSS feeds. Our experiment results show the robustness and efficiency of our approach.

Keywords Data warehouse · OLAP query log · Recommender system · RSS feeds

1 Introduction

The social media is becoming increasingly popular. Hence, the need to provide pertinent results on information retrieval proliferates. In this context, really simple syndication (RSS) is among the dedicated solutions for news monitoring according to various topics.

There is a large amount of previous work on processing RSS feeds from the user side through refining the RSS feeds according to the user requirements [6,7]. However, few works shed the light on analyzing the RSS feeds by decision makers to answer many questions, such as: Which is the most attention-grabbing RSS feeds topic? and How many RSS feeds are daily received? in order to undertake the compulsory actions.

Alongside, the set of RSS can be considered as a multidimensional data stored in a data warehouses (DW). In fact, due to the multiplicity of the analysis axis of RSS feeds, we opt for multidimensional modeling as a data warehouse in order to perform better-quality analysis. Indeed, according to Inmon [15], the data warehouse is an integrated collection of subject oriented, nonvolatile, historized, summarized and available data for analysis. Indeed, such data are organized according to analysis subjects in order to facilitate the extraction of relevant information. In such a context, OLAP operations can be performed in data warehouse to assist the decision maker [14]. The latter can ignore accurately the data warehouse schema, or the generated results of launched OLAP operations are distinct from the analysts expectations. Hence, the assistance of the analysis process arises through suggesting the MDX queries or OLAP fragments according to the analyst preferences.

In order to propose multidimensional recommendation, several approaches for OLAP recommender system have been proposed [3]. Among all these approaches, two main pools can be distinguished. The first pool approaches produce individual user recommendation based on graph models [16,17]. However, such approaches may lead to irrelevant recommendations when the number of analysts is so abundant. Hence, the second pool of approaches is introduced

✉ Eya Ben Ahmed
eya.benahmed@gmail.com

Nedia Araibi
nediaar@gmail.com

Wahiba Karaa Ben Abdessalem
wahiba.benabdessalem@isg.rnu.tn

1 High Institute of Management of Tunis, University of Tunis, Tunis, Tunisia

2 Miracl Laboratory, University of Sfax, Sfax, Tunisia

to suggest group recommendation through performing data mining techniques on users navigation histories to scrutinize the similarity between users [8–10,19].

Unfortunately, those approaches are not suitable for active data warehouse such as our RSS multidimensional modeling, and any eventual generated recommendation from classical strategies will be obsolete compared to our active context. Indeed, any extracted recommendation should have a feature that can integrate data changes while maintaining or scheduled cycle refreshes.

In this paper, we investigate another way aiming at introducing refined RSS recommendations based on a data warehouse perspective. Thus, we introduce a new RSS recommender system.

In summary, our major contributions are:

1. We formalize the problem of active data warehouse, and we introduce a new architecture for active data warehouse establishment;
2. We investigate the optimized RSS feeds modeling, and we define its multidimensional representation through introducing our active RSS warehouse functional architecture;
3. We sketch the active multidimensional modeling profile, and we propose an efficient devoted algorithm for drawing such a profile;
4. We shed the light on active recommendations issue, and we consolidate our proposal through introducing an IRORS recommender system;
5. We conduct a comprehensive experimental study to evaluate the performance of our system.

The remainder of the paper is organized as follows: Sect. 2 motivates our proposal through an inspiring example. Section 3 sketches a thorough study of the related work to the active data warehouse and multidimensional recommender system. Section 4 recalls formal preliminaries. Section 5 details our approach. Section 6 reports the experimental results, showing the soundness of our system. Finally, Sect. 7 concludes our paper and outlines avenues of future work.

2 Motivating example

We consider a motivating example from the social media area, namely the really simple syndication (RSS). Indeed, the RSS benefits [25] from a range of standard Web feed formats to bring out regularly updated information, such as news titles and, audio or video headlines. Therefore, an RSS feed incorporates relevant characteristics of news, namely the source or the channel providing such information, the publishing date and the news title.

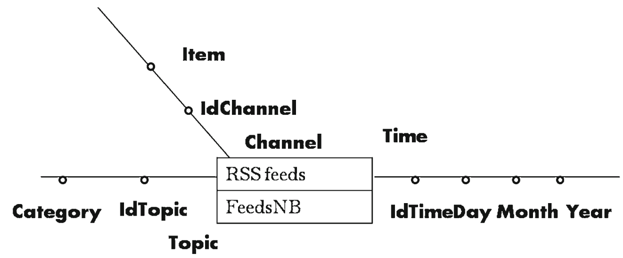


Fig. 1 A part of our RSS feeds data warehouse logical schema

Both publishers and users take advantages from RSS feeds. In fact, the latter facilitates for publishers to automatically organize data. Additionally, RSS feeds allow users to be up-to-date through getting timely updates from preferred Web sites avoiding to manually checking the Web site for novel content.

In this context, we propose an RSS feeds warehousing aiming to scrutinize the RSS feeds. The main goal behind new active RSS feeds modeling is to evaluate the progress of the news contents. Not only such carried out analysis of feeds allows an accurate identification of peaks periods and undertakes the required measures, but also it may stimulate feeds organized in subjects in order to permit a refined recommendation according to the users interests and avoid any overabundance of excessive flow.

A part of our logical data warehouse schema is depicted in Fig. 1. The adopted notations are similar to notations of [11]. Indeed, our schema includes a fact called RSS feeds which is measured using the number of feeds denoted as $FeedsNB$. It can be analyzed through various dimensions such as the channel, the topic and the time.

In this respect, according to Fig. 2, an analyst may launch a real-time query to discern in June 2015 the number of feeds related to the baccalaureate degree. Particularly, a more refined query can handle the numbers of feeds subject of declaration of results of baccalaureate degree in June 2015. Such a session may be used to establish the analyst profile interested on baccalaureate degree 2015, and we can recommend in real-time updated analysis of queries such as investigation of the success of baccalaureate degree in June 2015.

3 Related work

In this section, we present an overview of the previous work that has been done in active data warehouse and multidimensional recommender system.

3.1 Active data warehouse

Active data warehousing has appeared as an option to usual warehousing practices in order to converge the criti-

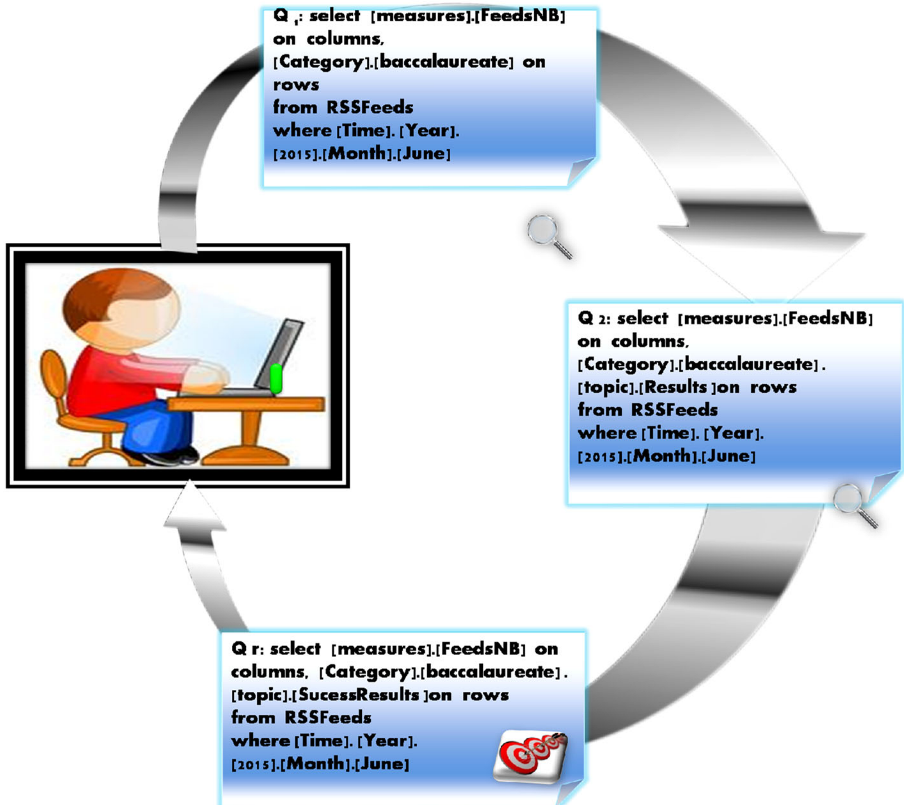


Fig. 2 An illustration of our recommender system process

cal request of applications for the newest information [21]. In this respect, in the literature, we distinguish two main trends of active data warehouse approaches. The first trend is principally based on the warehousing process analysis, and the second trend is essentially backboned on the ETL process.

In the first category, the usual data warehouse architecture is extended with analysis rules that imitate the activity of an analyst during decision making. Indeed [23], proposes a multidimensional analysis approach to define the analysis rules and to help analysts to change the values of given dimensions in order to improve the fact table.

However, in the second family of approaches [18,24], propose to follow updated tuples using messages holders of time indicators. The loading operation is performed continuously and must be done in real time admitting a consultation of data in parallel.

To the best of our knowledge, no method has incorporated both alimentation side and analysis side in the instauration of an active data warehouse. However, the update operation should impact both steps as new data will be considered during the alimentation step, but will affect the analysis stage in all cases. Hence, the need to introduce a new active data warehouse architecture that merges both steps is amplified.

3.2 Multidimensional recommender system

Recently, recommender systems have grabbed the attention of research society in data warehouse area. Indeed, Jerbi et al. [16,17] presented a framework for anticipatory OLAP recommendations in order to guide the user during his OLAP analysis through providing the imminent analysis step. To this end, a context-aware preference model has been introduced and applied to draw personalized anticipatory recommendations.

Along with the same preoccupation, Giacometti introduced a multidimensional recommender system taking benefits from the past querying experiences of analysts [8–10]. Its basic idea is to infer from the log of the OLAP server the activities of previous users to offer suitable recommendations to assist the user on his cube navigating.

Khemiri and bentayeb [19] presented a new approach to help the decision makers on constructing their analytical queries through investigating their past querying experiences. Indeed, the authors extracted frequent itemsets from the historical log files and used them to suggest pertinent recommendations, assuming that the significance of a recommended item wholly associated with its occurrence in the workload's queries.

More recently, Amo and Oliveira [2] presented a general framework for suggesting hybrid recommendations using preference mining and aggregation techniques. Indeed, they applied pairwise preference mining techniques to predict the favorite item for resolving the cold-start dilemma.

Moreover, Ben Ahmed et al. [4] introduced the semantic multidimensional group recommendations. Indeed, they presented the group profiling by means of an ontology on which they spread the analysts activities. Consequently, such profiling is exploited to develop semantic recommendations, namely fragments and complete queries. More recently, Aligon et al. [1] suggested recommending not single OLAP queries but whole OLAP sessions. First, the introduced recommender system recognized the most similar sessions with the current one. Then, the most significant subsessions derived and adapted to the current session.

Comparing the surveyed approaches in Table 1, we make out several analytical criteria, such as profiling source (behavior, external sources), user intervention (manual, automatic), profiling formulation (qualitative, quantitative), recommendation source (profile, external sources), recommendation strategy (language, method), type of recommendation (collaborative filtering [22], content-based), recommended object (fragment of query, complete query) and application domain (social media, movie, stock market, commercial).

It can be highlighted that all recommender systems are backboned on passive data warehouse. No consideration of real-time feature is involved. However, such a characteristic is fundamental in many domains such as social media where the speed of producing information is highly expedited and requires an instantaneous update of data warehouse.

Therefore, to the best of our knowledge, no strategy to suggest active recommendations is proposed in the literature. Thus, we introduce, in this paper, an innovative active data warehouse for RSS feeds modeling. From which, the users OLAP sessions are investigated to draw an efficient profiling of decision makers. Such a profiling is exploited to offer active recommendations of RSS feeds.

4 Preliminaries

In this section, we define the multidimensional model we will use to formalize our approach and illustrate it using our working example.

Definition 1 Multidimensional schema A multidimensional schema MS is a triple MS = (L, H, M) with:

- L is a finite set of levels, each level l ∈ L defined on a categorical domain Dom(l);
- H = h_1, ..., h_n is a restricted set of hierarchies, each characterized by: (*i*) a subset $L_i \subseteq$ L of levels (all L_i)'s are disjoint; (*ii*) a rollup total order > L_i of level (L_i);
- M is a limited set of measures, each defined on a numerical attribute.

Example 1 Figure 1 illustrates an example of multidimensional schema related to RSS feeds. Indeed, our schema MS= (L, H, M) is composed of:

- L is the set of levels related to all hierarchies, i.e., Year is a level of the time hierarchy;
- H is the set of hierarchies dedicated to all considered dimensions, namely Time, Topic and Channel;
- M is our measure, *i.e.*, FeedsNB.

Definition 2 OLAP query fragment
Given a multidimensional schema MS = (L, H, M), a query fragment f is either a level in L, a measure in M, or a simple boolean predicate involving a level and/or a measure.

Example 2 In our working example, a representative example of a query fragment is FeedsNB.

Definition 3 OLAP query
An OLAP query is a collection of OLAP query fragments with at least one level for each hierarchy in H and at least one measure in M.

Example 3 In our working example, an example of OLAP query expressed in MDX language incorporating diverse fragments (i.e., fragment as a measure and fragment as a level) is

SELECT [measures].[FeedsNB] **ON COLUMNS**, [Category].[baccalaureate] **ON ROWS FROM** RSSFeeds;

5 Our approach

In this section, we present our innovative concepts. Then, we describe our new recommender system.

5.1 Active preference aggregation model

Definition 4 Active preference P_i
Given a multidimensional schema MS of an active data warehouse, an active preference P_i is a fragment of OLAP query.

Example 4 A typical example of an active preference is FeedsNB in our working example.

Definition 5 Cadency of active preference
Let P_i be an active preference; the cadency of P_i in the context C, denoted as C(P_i), is defined as the number of occurrence of P_i in C.

Example 5 Given a context containing the active preference P_1= FeedsNB in three OLAP queries, the cadency of C(P_1= FeedsNB) is 3.

Table 1 Comparison of multidimensional recommender systems

Context		Approach	Profiling						Recommendation			
			Source		User intervention		Formulation		Source		Strategy	
			Behavior	External sources	Automatic	Manual	Quantitative	Qualitative	Profile	External sources	Language	Method
Multidimensional	Passive	Jerbi et al. [16,17]	X			X		X	X			X
		Khemeri and bentayab [19]	X		X		X		X			X
		Giaco-metti et al. [8–10]	X		X		X		X			X
		Ben Ahmed et al. [4]	X		X			X	X			X
		Amo et Oliveira [2]	X	X	X							
		Aligon et al. [1]	X		X		X		X			X
	Active	Our proposal	X		X		X		X			X
Multidimensional	Passive	Jerbi et al. [16,17]		X		X						X
		Khemeri and bentayab [19]	X			X						X
		Giaco-metti et al. [8–10]		X			X					X
		Ben Ahmed et al. [4]	X			X	X				X	
		Amo et Oliveira [2]			X					X		
		Aligon et al. [1]	X					X				
	Active	Our proposal		X		X			X			

Definition 6 Multidimensional active profile model
Let u be a user and P_i be an active preference; the multidimensional active profile model is denoted as MAP (u), and it is equal to the set of active preferences:
MAP $(u) = \cup P_i$.

5.2 Two-layered active data warehouse: RSS feeds case study

As shown in Fig. 3, our active data warehouse architecture enriches the classical data warehouse architecture through an innovative extension which consists on a novel layer to handle the update process in our data warehouse. In this subsection, we first introduce the classical modeling of our multidimensional data warehouse dedicated to RSS feeds in our case of study. Then, we detail our extended layer for insuring update operations.

5.2.1 Layer of RSS feeds modeling

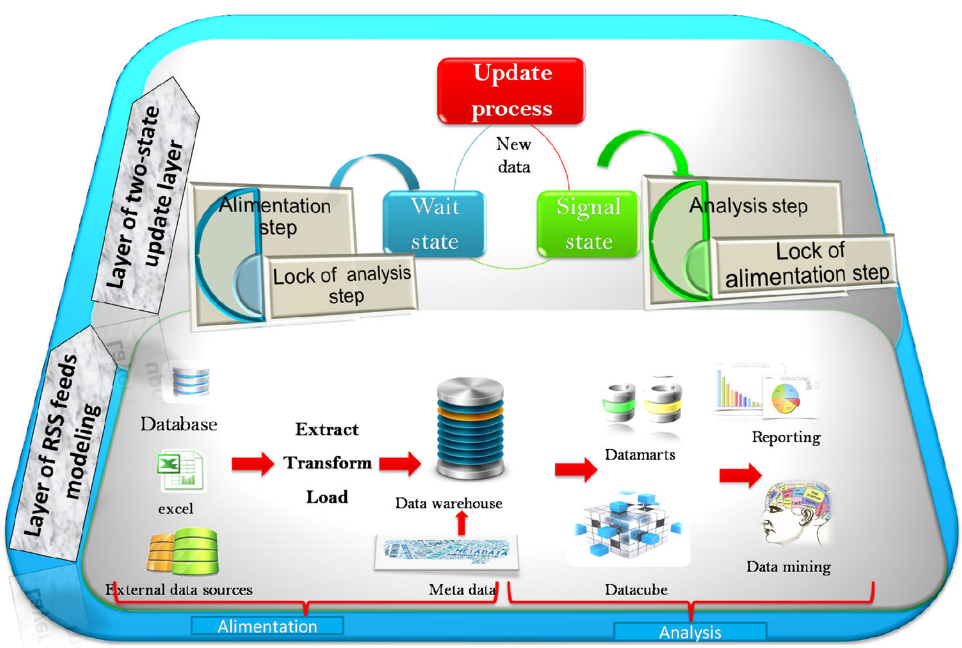
We propose to model the RSS feeds data as a multidimensional structure based on the STAR schema shown in Fig. 1. Indeed, the fact table RSS feeds includes the attribute FeedsNB that measures the number of RSS feeds. The latter can be analyzed from different perspectives, which are our dimensions. Thus, we introduce the Time dimension to report information about the date and the time when the RSS feed was delivered. Besides, the Topic dimension describes the topic of the RSS feed and its related category. Likewise, the Channel dimension depicts the metadata of the feed such as the item, the title and the associated URL. Moreover, hierarchies are used to aggregate the measure values. Therefore, we introduce a concept hierarchy for each dimension. For instance, Time [Day] Time [Month] Time [Year] is the hierarchy on the Time dimension, Topic [Name] Topic [Category] is the hierarchy on the Topic dimension, and Channel [Name] Channel [Item] is the hierarchy on the Channel dimension. Such aggregated hierarchies of dimensions can be used to sum up the measure values using operators like SUM, COUNT or AVG.

Aggregating measure values along the hierarchies of diverse dimensions (i.e., rollup) generates a multidimensional sight on data, which is recognized as data cube or cube. Deaggregating the measures of a cube to a lower dimension level (i.e., drilldown) provides a more detailed cube. Selecting the subset of a cubes cells that respect a given selection condition (i.e., slicing) also develops a more detailed cube. Such modeled layer must be maintained using our two-state update layer.

5.2.2 Two-state update layer

Regarding the update side, we introduce a two-state solution. Those states are the data warehouse positions during any maintenance operation. As soon as new data arises, the data warehouse mutates in the wait state where the alimentation of the new data is performed and any analysis operation is blocked in order to avoid any inconsistency in generated results. Once this maintenance is accomplished, the data warehouse moves to the signal state where analysis tasks

Fig. 3 Two-layered active data warehouse architecture



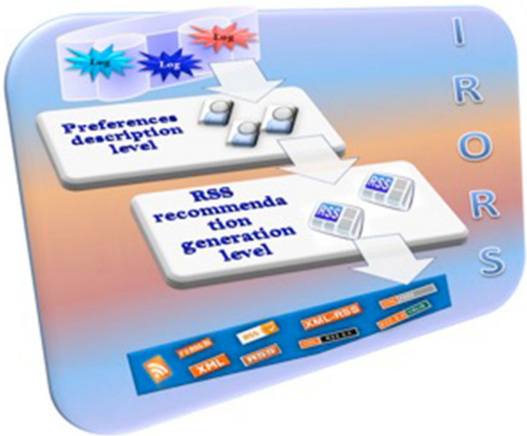


Fig. 4 IRORS architecture

can be performed at the moment that any update operation is blocked.

5.3 IRORS system architecture

Our new recommender system consists of two levels as shown in Fig. 4. The first level, known as multidimensional active profile model description level, handles the analyst preference modeling in our system. At the generation level, such modeled preferences are managed to draw appropriate RSS feeds recommendation using our new innovative strategy.

5.3.1 Multidimensional active profile model description

This level describes the multidimensional profile of active analyst through drawing his preferences. Indeed, all multidimensional active preferences are investigated from the OLAP log files and are used to draw the multidimensional active profile model of the decision maker. The key idea of our multidimensional active profile building is the computing of the cadency of each active preference. To do this, we introduce a new algorithm called ACTIF for automatic multidimensional active profile modeling. First, we scan the analysis context (Line 2), and we identify the category of each multidimensional active preference (Line 4). Then, we compute the cadency of each identified active preference in the scrutinized context (Line 5). Once the multidimensional active preferences Pi are collected in the multidimensional active profile MAP, the extracted profile is drawn (Line 8).

Algorithm 1 : ACTIF : multidimensional ACTIve proFile building algorithm
Input : Multidimensional schema MS, Analysis Context AC
Output : Multidimensional active profile MAP

```
1 Begin
2   Foreach query q in AC do
3     Foreach fragment f in q do
4       Identify the type of f
5       Pi → f
6       C(Pi) → C(Pi) +1
7       MAP → MAP U Pi
8   Return MAP
9 End
```

5.3.2 Generation of RSS recommendation level

At this level, based on identified multidimensional active preferences, our system generates the k candidate RSS recommendations. Finally, the display of generated recommendations is accomplished.

Extraction of RSS candidate recommendations Let X be an analyst who starts his query formulation. In addition, our algorithm ECAR takes as input the multidimensional active profile and the number of generated active recommendations k. First, ECAR captures the missing fragments on the current query MF (Line 4). For each missed fragment, it scans the multidimensional active preferences including in MAP (Line 5). When the type of the active preference corresponds to MF, it generates the current preference (Line 8). After that, it increments the flag variable A (Line 9). Such operation is stopped when our flag variable reaches the k parameter (Line 10).

Algorithm 2 : ECAR : ExtraCtion of Active Recommendation algorithm
Input : Multidimensional active profile MAP k : number of generated active recommendations
Output : Set of active recommendation R

```
1 Begin
2     a ← 0
3   Foreach query q do
4     MF ← Identify the type of missing fragment
5     Foreach Preference Pi in MAP do
6       Do
7         If Type(Pi)= MF then
8           R ← R U Pi
9           a ← a+1
10        Until (a=k)
11  Return R
12  End
```

Display of extracted RSS candidate recommendations After the extraction of RSS candidate recommendations, the results are displayed using matrix according to his requirements. Each cell of this matrix represents the related fact using the dedicated measure.

6 Experimental study

This section presents an experimental evaluation of the proposed system. First, we report our data description. Then, we discuss our experimental results.

6.1 Data description

All experiments have been performed on a system equipped with a 3-GHz CPU and 2 GB of main memory. In addition, they were conducted on real data warehouse built to support the analysis of RSS feeds. Undoubtedly, we collect 100 OLAP log files related to RSS feeds log files. Each log file contains 30 MDX queries. The whole size of analyzed files is about 3000 queries. For each query, we have identified the analyst preferences using his built profile.

In order to evaluate our proposal, we considered three quality measures broadly used, namely the recall, the precision and the F-measure [11,12]. We denote the number of recommended queries by card(R), the number of recommended significant queries by card(Rs) and the number of total significant recommended queries by card(RsT). Indeed, the precision describes the proportion of top results that are significant:

$$\text{Precision} = \frac{\text{Card}(Rs)}{\text{Card}(R)}. \tag{1}$$

However, the recall computes the proportion of all significant queries incorporated in the top results:

$$\text{Recall} = \frac{\text{Card}(RT)}{\text{Card}(RsT)}. \tag{2}$$

To assess the global accuracy, we apply the F-measure metric, computed as follows:

$$\text{F-Measure} = 2\frac{(\text{Precision} \times \text{recall})}{(\text{Precision} + \text{recall})}. \tag{3}$$

6.2 Results and discussion

This subsection describes the results of the experimental tests we performed. First, we report the scalability of our new recommender system. Then, to show the parameter sensitiveness and efficiency of the proposed system, we examine the correlation between the performance of IRORS and the cadency variation of multidimensional fragments. After that, we assess the pertinence of our IRORS system. Finally, the accuracy is discussed.

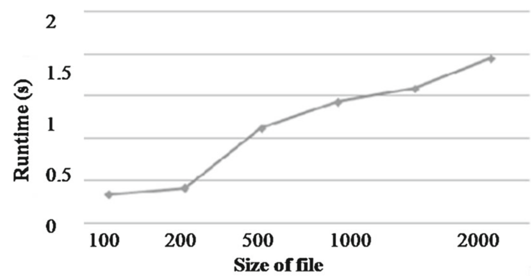


Fig. 5 Scalability of our IRORS system

6.2.1 Scalability analysis

In this subsection, we put the focus on the scalability of our system which describes the ability of our system to react facing the size variation of OLAP log files. To do so, we compare the runtime of our IRORS system versus the variation of OLAP log file size. The time of the recommendation increases regularly with the number of handled OLAP log queries. For example, if we handle 100 MDX queries, the required runtime of recommendations is 0.334 s. Regularly, an OLAP log file containing 300 queries needs 1.706 s in our recommendation process as depicted in Fig. 5. This can be explained by the increase in the size file engendering additional time for active profile building so that the recommendation process will exploit such a huge profile to generate active recommendations. Hence, more required time will be involved. These first results are intended to show the feasibility of our approach. They are encouraging as we do not observe notable changes in the recommender system runtime.

6.2.2 Performance analysis

To demonstrate the performance of our system, we scrutinized the impact of the cadencies variation. In fact, we have introduced randomly values of minimum confidence criteria denoted as MinC. Figure 6 sketches the statistics for the runtime obtained by IRORS versus the variation of the MinC values. In fact, we can point out that the MinC value influences the performance of our system, i.e., the runtime decreases as long as the MinC value increases. For example,

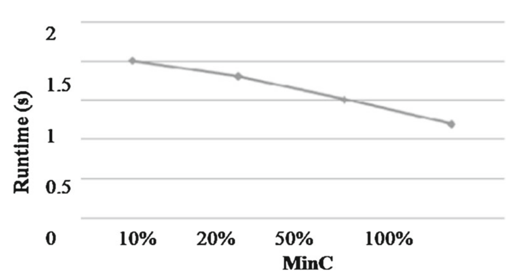


Fig. 6 Evaluation of IRORS performance

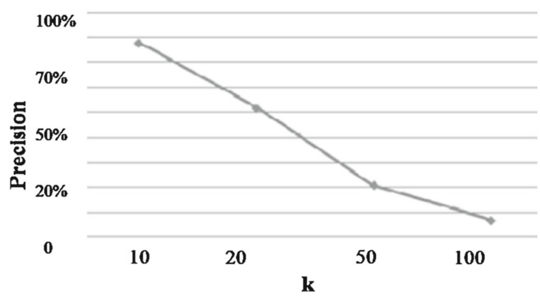


Fig. 7 Evaluation of the IRORS precision

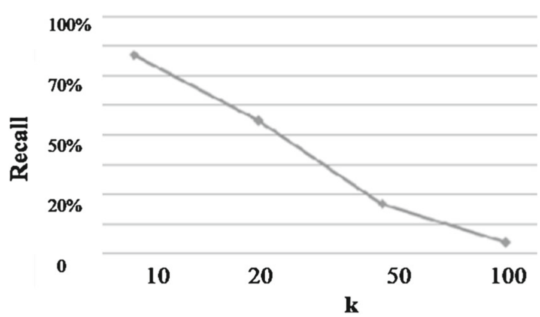


Fig. 8 Evaluation of the IRORS recall

a MinC value equal to 10 % engenders a runtime equal to 2 s. However, if we higher the MinC to 100 %, then we remark that the runtime decreases to 1.1 s. This can be explained by the fact that any rise of the MinC value leads to the vanishing of certain values, which became infrequent. This can hamper the performance of our recommender system.

6.2.3 Pertinence analysis

For each analyst, we extract k active recommendations dedicated to RSS feeds. We investigate the sensitiveness of this parameter on our IRORS system through varying its value from 5 to 100. Figure 7 shows the precision of our system according to the number of extracted recommendations.

In the resulting curve, presented in Fig. 7, we notice that the precision of our system gradually decreases when the number of generated RSS recommendations increases. This may be explained by the accuracy to recommend a restricted number of RSS recommendations.

Similarly, the recall, shown in Fig. 8, decreases when the number of suggested RSS feeds rises. It is important to mention that the recall deeply depends on the number of generated recommendations.

6.2.4 Accuracy analysis

The results in Fig. 9 show how the accuracy changes while the number of generated MDX query fragments k increases. When the number of candidate recommendations is 5, the best overall F-measure is achieved, meaning that the curve for precision clearly shows that our recommender system is well capable of delivering good recommendations when the low number of generated recommendations is input.

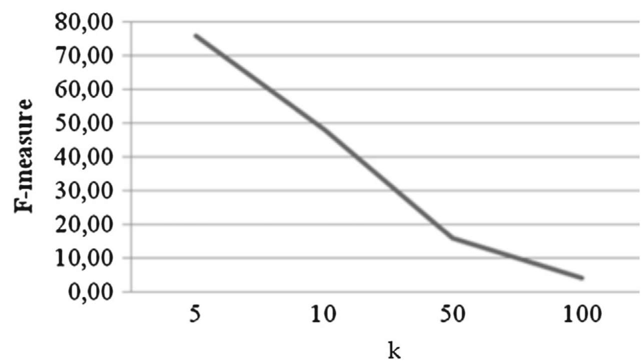


Fig. 9 Evaluation of IRORS accuracy

7 Conclusion

In this paper, we proposed a new active data warehouse architecture involved by two dedicated layers. Based on this, a new recommender system to assist the OLAP analysis of RSS feeds is proposed. Such a proposal is managed by three-step process: (i) modeling of the active user profile from OLAP log files using our ACTIF algorithm; (ii) extraction of candidate recommendations from the identified active preferences using our new algorithm ECAR; and (iii) display of relevant RSS recommendations. To evaluate our approach, we conducted several experiments that have highlighted the performance of our recommender system.

As future work, we finally plan to: (1) evaluate the recommender system quality that determinates the recommendation reliability, and (2) integrate the uncertainty aspect to deal with our imperfect context.

References

1. Aligon, J., Gallinucci, E., Golfarelli, M., Marcel, P., Rizzi, S.: A collaborative ltering approach for recommending OLAP sessions. Decis. Support Syst. **69**, 20–30 (2015)
2. Amo, S., Oliveira, C.G.: Towards a Tunable Framework for Recommendation Systems Based on Pairwise Preference Mining Algorithms, 27th Canadian Conference on Artificial Intelligence, Canadian AI 2014, Montral, QC, Canada (2014)

3. Ben, Ahmed E., Nabli, A., Gargouri, F.: A survey of user-centric data warehouses: from personalization to recommendation. Int. J. Database Manag. Syst. **3**(2), 59–71 (2011)
4. Ben Ahmed, E., Tebourski, W., Ben Abdesselam, W., Gargouri, F.: SMART: semantic multidimensionAl gRroup recommendations. Multimed. Tools Appl. **74**(23), 10419–10437 (2014)
5. Burke, R.: Hybrid recommender systems : survey and experiments. User Model. User-Adap. Inter. **12**(4), 331–370 (2002)
6. Creus, J., Amann, B., Travers, N., Vodislav, D.: RoSeS: a continuous content-based query engine for RSS feeds. In: Proceedings of 21th International Conference on Database and Expert Systems Applications DEXA11, LNCS, Toulouse, pp. 203–218 (2011)
7. Creus, J., Amann, B., Travers, N., Vodislav, D.: Un agrgateur de ux rss avanc, 26 me Journes Bases de Donnes Avances (2010)
8. Giacometti, A., Marcel, P., Negre, E., Soulet, A.: Query recommendations for OLAP discovery-driven analysis. Int. J. Data Warehouse. Min. **7**(2), 1–25 (2011)
9. Giacometti, A., Marcel, P., Negre, E., Soulet, A. : Query recommendations for OLAP discovery-driven analysis. In: International Workshop on Data Warehousing and OLAP (DOLAP), pp. 81–88 (2009)
10. Giacometti, A., Marcel, P., Negre, E. : Recommending multidimensional queries. In: International Conference on Data Warehousing and Knowledge Discovery (DaWaK), pp. 453–466 (2009)
11. Golfarelli, M.: From user requirements to conceptual design in data warehouse design—a survey. Data Warehous. Design Adv. Eng. Appl. Methods Complex Constr. **23**(1), 1–16 (2008)
12. Gunawardana, A., Shani, G.: A survey of accuracy evaluation metrics of recommendation tasks. J. Mach. Learn. Res. **10**, 2935–2962 (2009)
13. Ge, M., Delgado-Battenfeld, C., Jannach, D.: Beyond accuracy: evaluating recommender systems by coverage and serendipity, pp. 257–260. Barcelona, Spain, Proc. RecSys (2010)
14. Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques, Series in Data Management Systems. Morgan Kaufmann Publishers, Burlington(2011)
15. Inmon, W.B.: Building the Data Warehouse, 4th edn. Wiley, New Delhi, India (2005)
16. Jerbi, H., Ravat, F., Teste, O., Zuruh, G.: Applying Recommendation Technology in OLAP Systems. In: ICEIS Conference Proceedings, pp. 220–233 (2009)
17. Jerbi, H., Ravat, F., Teste, O., Zuruh, G.: Preference-based recommendations for OLAP analysis. In: DaWaK Conference Proceedings, pp. 467–478 (2009)
18. Karakasidis, A., Vassiliadis, P., Pitoura, E.: ETL Queues for Active Data Warehousing. In: IQIS05: Proceedings of the 2nd International Workshop on Information Quality in Information Systems. ACM Press, New York, NY, pp. 28–39(2005)
19. Khemiri, R., Bentayeb, F.: Interactive query recommendation assistant. In: The 23rd International Workshop on Database and Expert Systems Applications (DEXA), pp. 93–97 (2012)
20. Lops, P., de Gemmis, M., Semeraro, G.: Content-based recommender systems: state of the art and trends. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook pp. 73–105. Springer, USA (2011)
21. Mohania, M., Nambiar, U., Schrefl, M., Vincent, M.: Active and Real-Time Data Warehousing. Encyclopedia of Database Systems. Springer, USA (2009)
22. Su, X., Khoshgoftaar, T.M.: A survey of collaborative ltering techniques. Adv. Artif. Intell. **2009**, 421425 (2009)
23. Thalhammer, T., Schre, M., Mohania, M.K.: Active data warehouses: complementing OLAP with analysis rules. Data Knowl. Eng. **39**(3), 241–269 (2001)
24. Tho, M.N., Tjoa, A.M.: Zero-latency data warehousing for heterogeneous data sources and continuous data streams. Services Computing, pp. 357–365 (2004)
25. Travers, N., Hmedeh, Z., Vouzoukidou, N., Mouza, C., Christophide, C., Scholl, M.: RSS feeds behavior analysis, structure and vocabulary. Int. J. Web Inf. Syst. **10**(3), 291–320 (2014)