To study and analyze Information Filtering approaches for Emergency responders during Crisis Management[1]

Vedna Sharma
Junior Research Fellow
Institute of Technology Management
(Defense Research & Development Organization)
Landour Cantt. Mussoorie Uttarakhand, India

Abstract

Information filtering is a function to select useful information for the user among a large amount of information. Among Organizations major distinctive feature of information filtering is transition from the classical activity of Information Retrieval to organized information Filtering, then conversion of filtered information into Knowledge to help decision making. This function has become very important as network technology develops rapidly and to making fast decision with overcome cost of time with the increasing value of time among organization. This paper focuses on information filtering for emergency management. When a large-scale disaster happens, the problem of information flood can be very serious because a great deal of information occurs in a short time and is sent to a person or an organization that is responsible for managing the situation. Information filtering systems filter data items by correlating vector of terms that represents data items e.g. documents, emails.

In previous years, neural networks have yielded immense success on speech recognition, computer vision and natural language processing. However, the exploration of deep neural networks on Information Filtering systems has received relatively less scrutiny. Many IF Systems have been developed in previous years for various application domains. Some examples of filtering applications are searching personal mails based on personal profiles or newsgroup filter for groups or individuals. In this paper we emphasis to study and analysis techniques for acquiring knowledge of user for information filtering, various concepts, approaches and components in IF System.

Keywords: Information Filtering, Emergency Management, user profile, Deep Learning, Information filtering Types

[1] How to cite this paper: Sharma, V. (2018). To study and analyze Information Filtering approaches for Emergency Responders during Crisis Management; *PM World Journal*, Vol. VII, Issue 6 - June.

1. Introduction

Information filtering (IF) is one of the methods that are rapidly evolving to manage large information flows. It deals with the delivery of information that the user is likely to find interesting or useful. Because knowledge represents an important aspect for an organization and a way to compete more efficiently, strategic information systems. In front of the critical proliferation of electronic information and the underlying difficulty to manage this information in a relevant way, the usual answer is to reduce drastically the volume of documents available to the end users using abstracting or filtering process.

Information filtering combines tools from the field of Artificial Intelligence (AI), such as intelligent agents or software robots, with information retrieval approaches, indexing and retrieving of content [3].

Keyword selection for the specification of user profiles or queries is important and complex task sometimes. In this paper we study and analyze the concepts and discuss the flow chart and working methodology of information Filtering System which uses all meaningful words in document space as input data.

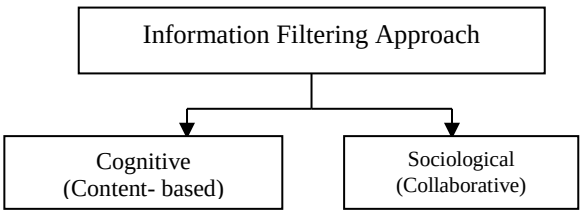
The rest of this paper is structured as follows: Section 2 review mainly used approaches in Information Filtering and its features. Section 3 describes possible techniques of acquiring knowledge about users for information filtering. Section 4 provides brief challenges of information filtering. Section 5 represents component and methodology of IF System. Section 6 provides Information about the application of Information Filtering System in Emergency Management.

2. Approaches used for Information Filtering

The functionality of filtering systems is to successfully separate relevant and non- relevant documents in an incoming stream of textual information. When the delivered information comes in the form of suggestions an information filtering system is called a recommender system.

Information Filtering consist mainly two approaches which are:

Content-based Filtering (Cognitive)
Collaborative Filtering (Sociological)



Information filtering (IF) is one of the methods that is rapidly evolving to manage large information flows. The aim of IF is to expose users to only information that is relevant to them. Information filtering (IF) covers a broad range of domains, technologies and method involved in the process of exposing user to the information they need. Due to the timeliness of information and the dynamic and ever changing environment these filters can't be static, as the individual's needs change [10].

Information filtering (IF) systems are mainly consist features like following:

- Information Filtering System are applicable for unstructured or semi-structured data (e.g. documents, e-mail, messages)
- They handle large amounts of data
- Information Filtering System deal primarily with textual data
- They are based on user profiles
- Their objective is to remove irrelevant data from incoming streams of data items.

2.1 Cognitive (Content- based) Filtering Approach

Content-based filtering also referred to as cognitive filtering, recommends items based on a comparison between the content of the items and a user profile. The content of each item is represented as a set of descriptors or terms, typically the words that occur in a document. The user profile is represented with the same terms and built up by analyzing the content of items which have been seen by the user. Content-based method, also referred as content-based filtering, assumes that the user show the same particular behavior under the same circumstances. The information source that content-based filtering systems are mostly used with is text documents. A standard approach for term parsing selects single words from documents.

The learning methods applied to content-based filtering try to find the most relevant documents based on the user's behavior in the past. Such approach however restricts the user to documents similar to those already seen. It is known as over- specialization problem as stated before the interest of a user are rarely static but change overtime. Instead of adapting to the user's interests after the system has received feedback one could try to predict a user's interests in the future and recommend documents that contain information that is entirely new to the user [7].

2.2 Sociological (Collaborative) Filtering Approach

The goal of collaborative filtering (CF) is to predict the preferences of one user, referred to as the active user, based on the preferences of a group of users [8]. Collaborative filtering also referred to as social filtering, that filters information by using the recommendations of other people. It is based on the idea that people who agreed in their evaluation of certain items in the past are likely to agree again in the future. Unlike content based method, collaborative method

ignores the item's content and does recommendation of the items only based on the similar users' item ratings. In the more general sense, collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources etc. Applications of collaborative filtering typically involve very large data sets. Collaborative filtering systems infer knowledge about the user; from his/her similarity to other users and the filtering process is based on those similar user recommendations.

Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc.

Collaborative filtering system involves the computation of similarity between user interests. Similarity between the users interests are calculated using various methods. The system collects the ratings of each item from different users explicitly and calculates the similarity between the ratings of the users. The ratings can be explicit on a numeric scale or implicit such as shopping, links clicks and mouse movement. Then, the users are grouped based on the calculated similarity measures and future items are recommended to the user based on the recommendation of other users in the group [5].

For example: Consider a group of users X1, X2....Xn and items Y1, Y2....Ym. The table below shows the rating given by the users on different items.

	Y1	**Y2**	**Y3**	**…..**	**Ym**
X1	1	4	4		5
X2	2	4	3		3
X3	1	4			5
X4	2	4	3		
					4
….					
Xn	3	4	1		4

Table 1: Rating given by the users on different items

If similarity rating between the user X1 and X3 is high, then user X1 and X3 can be grouped and new items will be recommended to each user based on the other user's interest. Here, item Y3 will be recommended to the user X3, as a new item based on the high rating given by the other user in the group X1. Similarly, item Ym will be recommended to User X4 based on the rating of other user X2.There are two types of collaborative filtering systems are available: memory based and model based.

Memory based collaborative filtering systems works on the neighborhood principles to recommend items and model based collaborative systems uses models in recommending items by estimating models based on the ratings.

3. Knowledge acquiring Techniques for an Information Filtering System

3.1 User Interrogation

It is the most popular technique of acquiring knowledge about user. Systems utilizing this method usually require their users to fill out a form describing their areas of interest or other relevant parameters.

Some filtering systems provide the user with a predefined set of policies from which user may choose the most suitable profile. Other systems provide the user with a set of terms that represent each domain, from which user can construct a personal profile. This method prevents the semantic confusion observed in systems where users have the freedom to choose their own terms. There are rule-based filtering systems that utilize the user interrogation method to define rules of information filtering. These systems usually provide a rules editor that guides the user in the task of rules definition [7].

3.2 Recording User Behavior

This is an implicit approach that does not require active user involvement in the knowledge acquisition task. The user's reaction to each incoming data item is recorded, to learn the actual relevancy of the data item to the user. User behavior like whether the user saves, discards, prints or forward data items can serve as an indication of interest.

In the system learns the user profile by passively observing the hyperlinks clicked on and those passed over and by measuring user mouse and scrolling activity in addition to user browsing activity.

3.3 Document Space

This method creates a field of documents that the user has previously. This method is for acquiring knowledge about user's lies between the explicit and the implicit approaches as it require minimum user involvement. This method creates a field of documents that the user has previously judged as relevant. Any new incoming document is tested for its similarity to the documents existing in that space. Usenet filtering system, Newsweeder is an example of a system that uses this technique. In this system the users are required to rate documents for their relevance.

3.4 Stereotypic Inference

Stereotypes have a powerful role in information filtering. It can be used to build an initial user profile or to add knowledge to a user profile from the user's stereotypic belonging. It is also combined the explicit and implicit approaches. The Stereotype is one of the common elements in many users modeling works. It captures default information about groups of people. Stereotypes are also used in User Modeling as a User modeling tool as a source of initial default information to model the users when no other information is available.

4. Challenges of Information Filtering

User profile has to specify current and future documents correctly. The quality of user profiles is a key to making a filtering system work. From the user's point of view, there are two potential problems. One is if a large proportion of the items that the system sends to a user are irrelevant and then the system becomes more of an annoyance than help. Conversely, if the system fails to provide the user with enough relevant information, then the benefit of information filtering delivery is largely lost, because the user will still have to actively hunt for information. Explication of long-dated information needs in the shape of a user profile is a difficult job, as the user needs and interests depend on many parameters, including the personal & behavioral characteristics. Some of the parameters are not easily translated into system components (user's mood, work load, financial position, etc.).

Total utility of an Information Filtering system is maximal, if effort in programming the user profile is low and the quota of correct classified documents is high. If the user profile is of high volume, make filtering system over burden.

Humans are often not able to explain in general how they evaluate the relevance (class affiliation) of a document, i.e. the user may not be able to explain the basis on which he has given the relevance feedback to filtering system. Privacy is a important issue in maintaining the user profiles [6].

5. Components of Information Filtering System

An Information Filtering System contains four main components to analyze data items and to represent them efficiently and to acquire knowledge on the users and to represent it in user models. Four main components includes in Information Filtering System are:

1. User Model Component
2. Data Analyzer Component
3. Filtering Component
4. Learning Component

User Model Component: The User Model Component implicitly gathers information about the users and their information needs and constructs user profiles. The user model will be the input to the filtering component.

Data Analyzer Component: The data analyzer component obtains data items e.g. documents from information providers. The data items are analyzed and represented in index format. It provides input to the filtering component.

Filtering Component: The filtering component is main part of the IF system. It matches the user profile with the represented data items and identifies the data item is relevant to the user or not. Sometime filtering process is applied to single data item e.g. email message and some time it is applied on set of data items e.g. pool of documents. User evaluation enables further feedback to be provided to the learning components as shown in above flow chart.

Learning Component: The learning component needed to improve further filtering. Filtering Systems must include a learning process that detect shifts in user interests and update the user model. The learning component in statistical filtering systems utilizes user feedback to update the user profile.

6. Working Methodology of Information Filtering System

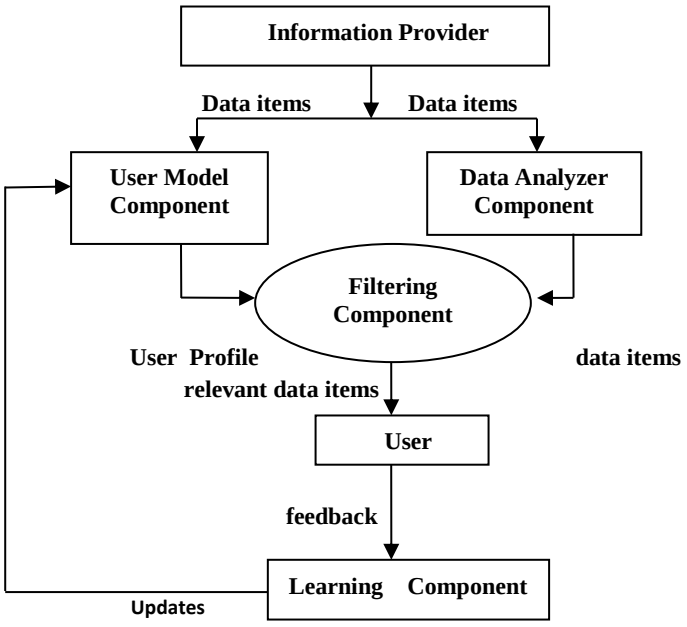


Fig 1. Working System of Information Filtering

In above flow chart we have discussed the working system of Information Filter step by step. First of all data items are gathered from information providers and provide as input to the user

model component and data analyzer component. Data analyzer component analyzed the data items and represent the data items in index form.

On the other hand The User Model Component gathers information about the users and their information needs and constructs user profiles. Output of both Data Analyzer and user model component is provided user profile and updated data items as input to the Filtering Component. Filtering System matches the user profile with the represented data items and identify the data item is relevant to the user or not. Relevant user data items filtered by Information Filtering system provided to the user [7].

After that the learning component in statistical filtering systems utilizes user feedback to update the user profile. Sometimes it become difficult in modeling user to changes in their information needs therefore filtering system must include in learning system that detects changes in user interests and update the user model by updating or changing or canceling the existing knowledge about the user. Otherwise inaccuracies present in profile of user that affect the information filtering system.

7. Applications of Information Filtering in Emergency situations

- Information Filtering system provided domain specific knowledge to improve decision-making quality in a timely manner and the established the foundations for decision improvement.

- In Security Management information filtering used to achieved data access control by implementing "need to know" policy for information security.

- In Contact Management it provides centralized repository for permanent and temporary response personnel or group.

- For Preplanning Manages organizations and personnel needs information filtering system helps for provide predesigned plans for emergency mitigation and recovery.

- Information Filtering System helps in Incident Action Plan by ensure clear management objectives.

8. Conclusion

At emergency situation or in the disaster environment the accurate information or exact knowledge about the users who can handle that situation is requiring essentially.

Efficient information retrieval requires information filtering and search adaptation to the user's current needs, interests, knowledge level, etc. From early days of SDI (selective Dissemination of information) to current modern information retrieval, information filtering has undergone a tremendous change. In this paper I have tried to focus on methods of Information Filtering. I have not covered all aspect of the topic like the technical details behind the methods because it is behind the scope of this article. My main aim of this paper is to pinpoint the highlights of Information Filtering, and to draw a comprehensive picture of methodology behind Information Filtering.

References

[1] arxiv 1708.05031[cs.IR] Neural Collaborative Filtering 26 Aug 2017

[2] F.O. Isinkaye a, *, Y.O. Folajimi b , B.A. Ojokoh c, Recommendation systems: Principles, methods and evaluation, International Journal of Computer Applications, 3, December 2014.

[3] Zvi Boger, Nuclear Research Center - Negev, and Optimal – Industrial Neural Systems Ltd., Beer- Sheva, Information Filtering and Automatic Keyword Identification by Artificial Neural Networks Israel 28 May,2014

[4] Ayse Cufoglu Anglia Ruskin University Department of Computing and Technology Cambridge User profiling - A short Review Cambridge, December 2014

[5] Renganathan V, Babu AN, Sarbadhikari SN (2013) A Tutorial on Information Filtering Concepts and Methods for Bio-medical Searching. J Health Med Informat 4:131. doi: 10.4172/2157-7420.1000131 September 09, 2013

[6] Vinaitheerthan Renganathan1 *, Ajit N Babu2 and Suptendra Nath Sarbadhikari3 A Tutorial on Information Filtering Concepts and Methods for Bio-medical Searching, 2013

[7] Information Filtering: Overview of issues, Research and system Department of Information systems Engineering 5 December, 2000

[8] Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and Model-Based Approach , Uncertainty in Artificial Intelligence proceeding, 2000

[9]O'Donovan J., Dunnion J. (2004) Evaluating Information Filtering Techniques in an Adaptive Recommender System. In: De Bra P.M.E., Nejdl W. (eds) Adaptive Hypermedia and Adaptive Web-Based Systems. AH 2004. Lecture Notes in Computer Science, vol 3137. Springer, Berlin, Heidelberg

[10] Eberhagen, N. (1993) Information filtering. In: Sundsvall 42: ADB i verksamhetens tjänst Sundsvall, Sweden: Sundsvall 42 and The Swedish Information Processing Society

[11] Jacob Palme Department of Computer and Systems Sciences, http://www.dsv.su.se/~jpalme/select/information-filtering.pdf

[12] Callan J (2000) System and method for filtering a document stream. Google Patents US.

[13] O'Riordan, Adrian & Sorensen, Humphrey. (1995). An Intelligent Agent for High-Precision Text Filtering.. 205-211. 10.1145/221270.221569.

[14] Content-boosted collaborative filtering for improved recommendations P Melville, RJ Mooney, R Nagarajan - Aaai/iaai, 2002

[15] Sheth B, Maes P (1993) Evolving agents for personalized information filtering. Artificial Intelligence for Applications. Proceedings Ninth Conference, IEEE.

About the Author

Vedna Sharma

Mussoorie Uttarakhand, India

Ms Vedna Sharma received M.tech & B.tech degrees in Computer Science from Himachal Pradesh Technical University & HPU Shimla respectively. She is currently working as Junior Research Fellow (JRF) at the Institute of Technology Management, Mussoorie Ministry of Defense (DRDO), Landour Cantt. Mussoorie Uttarakhand, India. She can be contacted at sharma.vedna@gmail.com