PRISENIT – A Probabilistic Search Recommendation Algorithm to Improve Search Efficiency for Network Intelligence and Troubleshooting

Mikel Zuzuarregui*, Enda Fallon#, MingXue Wang*, John Keeney*, Paul Jacob*

*Network Management Laboratory
Ericsson Software Campus
Athlone, Ireland
{mikel.zuzuarregui, mingxue.wang, john.keeney}@ericsson.com

#Software Research Institute
Athlone Institute of Technology
Athlone, Ireland
{efallon, pjacob}@ait.ie

Abstract—When searching for data in a telecommunications network management application, large search result sets are common. In order to refine the results to retrieve useful information existing systems normally require additional user intervention such as appending or removing a search keyword, adding a filter, grouping results, etc.. This work proposes a Probabilistic Search Recommendation Algorithm to Improve Search Efficiency for Network Intelligence and Troubleshooting (PRISENIT). PRISENIT is a query-based recommendation algorithm intended to improve search efficiency and usability for telecommunication system management. PRISENIT is an extension of an item-based collaborative filtering algorithm. It uses correlation-based similarity and users’ implicit feedback in order to improve search efficiency. It learns from previous experiences in order to optimize decision-making. Currently there exists no known query-based recommender adaptation mechanism for network management. Existing search engines use previous user searches to make a suggestion based on the keyword. PRISENIT not only considers search terms, it also considers the influence of filters and features in order to make network searches more efficient as it removes the necessity for users to manually choose search features or search filters. Experimental results show that PRISENIT can improve user experience in a telecommunications management environment.

Keywords - recommender system; item-based collaborative filtering; query-based recommender; network management

I. INTRODUCTION

With the ever increasing volume of telecommunications data (end user data, performance metrics, and alarms), big data and information retrieval (search) technologies have gained attention by telecom equipment vendors [1][2]. In network management systems traditional SQL-based data-retrieval techniques are mainly used. Such systems however are performance limited for unstructured telecom data such as network node logs, alarm text and trouble ticket messages. Increasingly enterprise search technologies are capable of handling both structured data from corporate databases and unstructured data sources and have potential for management analysis. The characteristics of telecommunications network data however, volume, complexity and differing lifespans create challenges when developing network analysis solutions.

This work particularly focuses on how a search interface can be improved to assist network troubleshooting. PRISENIT can learn from previous experiences and interact with human decision techniques. In the case of system error, it can implement autonomic system reconfiguration. Unlike existing search approaches, PRISENIT considers not only query terms (keywords) but also search filters and presentation features. In other commercial solutions it is possible for users to manually store and maintain interaction templates. Within the PRISENIT approach common searches are automatically learned and recommended. PRISENIT uses correlation-based similarity and combines implicit feedback from the user in order to improve search efficiency for network intelligence and troubleshooting. Such learning and query suggestion is not currently available for telecommunication or network management domains. Data Analytics approaches (e.g. OLAP tools) also support the user to define custom search-based reports (custom filters and features to present results). However, these reports are not automatically learned, tuned or recommended to the user. As the searched data changes, or preferred ways of interacting with the system changes, our system learns and recommends new queries (terms, filters and features). PRISENIT speeds up network management system interactions for common or routine management tasks, and also learns best practices for troubleshooting, reducing skill requirements and cost in solving management issues.

In summary, the key problem is that users usually require multiple additional iterative search actions or steps after an initial query, such as appending or removing a keyword, adding a filter to find results, and then selecting presentation features to appropriately present the results. This makes search-based management complex to use and tedious and time consuming for users, while also expending system resources performing searches that are unnecessary.

We empirically evaluate PRISENIT with data from experimental evaluations undertaken at the Ericsson research lab. The system was evaluated using Ericsson network management engineers. Results illustrate the enhanced search interactivity of the PRISENIT approach.

The remainder of this paper is organized as follows. Section II will introduce related work about how to personalize search engine interfaces. Section III shows the architecture and methodology of the recommender system. Section IV will show the system application. In section V, we discuss the usability of the recommender system, and point out how a prototype implementation was developed and operated, before concluding in section VI.
II. RELATED WORK

A. Search engine

The explosive growth of Web search services has triggered an increasing demand for web personalization systems in order to improve end user Web browsing experience. In the recent years such systems have become ubiquitous. Typically such systems take advantage of data about the users’ past behaviors or usage. Web personalization can typically be categorized in three categories: Web Utilization Miner (WUM), user profiling and recommender systems.

Web mining techniques have been analyzed for several years. In [3] the authors investigate the importance of a suitable data set. They introduce a novel model for improving web-mining using semantic web and synthesis web. In [4] the impact of semantic web to improve web mining experience is examined by presenting a weight estimation process with span time, request count and access sequence details.

A number of approaches have been investigated in the area of user profiling. In [5] the authors provide a framework and findings in mining web-navigation from website log files, involving user profiles and external data describing the ontology of the web content. In [6] the authors propose description-logic based semantic user modelling for spatial web personalization. Recently, ontologies have been applied to obtain user profiles in [7].

B. Web personalization

The challenge of Web personalization and usability can be addressed by recommender system approaches. Recommendation systems use information filtering to assist a user in a decision-making process by giving some suggestion that may benefit the user based on previous experience of that user and similar users. The suggestion could be about products, information or services by learning user preferences. A number of studies focus on recommendation systems and web personalization, e.g. using ontologies [8] or sequential web access patterns [9].

C. Recommender system

Recommendation systems first appeared circa 1997 [10]. In general, all the recommender systems are based on the analysis of item attributes, user attributes or user behavior. A recommendation system must predict the utility of a candidate item in comparison to alternatives.

Collaborating Filtering was proposed by Rich [11] and has been applied in many commercial web sites. It consists of n users and m items connected by user ratings of items. Ratings can be explicit (e.g. based on feedback) or implicit (e.g. based on uptake). Implicit feedback is more difficult to determine however it can usually provide more information and context than explicit feedback.

In [12] a query language for network search is introduced. The search space is represented with objects as a set of attribute-value pairs, additionally showing that the method allows distributed execution.

Previously in [13] we investigated the applicability of recommender systems to assist Network Operations Centre operators for network management systems. In this paper we create a monitoring and trouble-shooting system for network managers.

PRISENIT is unique in that it focuses on using a search engine for network management to improve network intelligence and troubleshooting by recommending improved query-based searches. It shows the user how to change the search parameters to reach a possible better solution, as well helping to efficiently show more meaningful information for particular use cases.

III. PRISENIT ALGORITHM

A. Transaction log

A transaction log is a log file of the communications between a system and the user of that system. It has two main parts: a) Search query: A search query part represents the complete user search requirements; b) Rating: the rating part is the feedback of result of the search

The query part is a vector $V = [\nu_1, \nu_2, \nu_3]$ that can be further divided into three components; terms, filters, and features. These three parts cover the user requirement of different aspects of a search.

$\nu_1$ is a user query which is composed of a number of keywords or terms. Terms are a series of characters that the users type in the search box within a query separated by space.

$\nu_2$ is a set of filters applied on the search results such as a specified time range or restrictions on the data source, etc. $\nu_3$ describes presentation features with different settings to describe how the result is presented. For example, the result is sorted by time or by ranking score, the result is faceted by a node or/and severity, etc.

The rating part $\nu_4$ is a user rating of the results returned from the query part. It is measured by user interaction / behavior with result (i.e. time spent interpreting/using the results, using scrolling, click-through, and so on).

B. The recommendation calculation

The search recommendation is based on past user search experience. PRISENIT calculates a recommendation score for records in the translation log for a target user input query. Fig. 1 illustrates the mechanism utilized by PRISENIT when calculating recommendation scores.

Within PRISENIT, the computation of similarity of a query is composed of the amalgamation of the similarity of each component of the query using a pluggable correlation mechanism. Taking the term component $\nu_1$ as an example, the computation of similarity value of terms of a target query with terms of a comparing query is undertaken using the Pearson correlation coefficient $\text{sim}_T.\nu_1$. Similarly, we compute the similarity between the filter and feature element using the similar $\text{sim}_F.\nu_2$ and $\text{sim}_F.\nu_3$ functions. In contrast, a set of common filters or common features is used for the measure. The selection of the Pearson correlation within PRISENIT is based on [14][15].
The search rating is based on the end user interaction with the suggested results. A user action $u$ is a measure of some type of user interaction; mouse-overs, click-throughs, viewing different results pages, scrolling windows and time spent on the page, etc.

$$\sum_{i=1}^{n} a_i w_i.$$  \hspace{1cm} (1)

$w_i$ is a weight factor applied to each type of user interaction, $\sum_{i=1}^{n} = 1$. A higher weight is applied to user actions which apply directly to results; clicking on results or comparing actions such as time spent on the page.

In order to consider user feedback on search recommendations, a score is calculated based on both the similarity of the query and the rating.

$$\sum_{i=1}^{n} \text{sim}_p = r.$$  \hspace{1cm} (1)

Finally, the queries with the top recommendation scores are used to create the search recommendation. The recommended search is a combination of each component of the search query. It considers all terms, filters and features in order to create a new query.

$$V_{|2|} \cup \bigcup_{i=1}^{n} V_i.$$  \hspace{1cm} (1)

The union is over the $k$ queries with the highest scores.

Table 1 illustrates the queries that are stored in the database. PRISENIT uses those queries as the basis for comparison. Table 2 illustrates the operation of the PRISENIT recommendation algorithm using the queries from Table 1. In this example, the queries with the most similarity are 2, 3 and 5.

---

**Algorithm 1** Recommendation calculation

**Input:** Q String with the query

**Output:** String with the recommendation

**Initialization** finalScores;

1. split the query into terms, filters and features;
2. queries = retrieveSetsOfQueriesWithRatings();
3. for each queries do

   4. $w_{|p|} = \text{Correlation}(\text{terms}, \text{queries}_{|1|})$;
   5. $w_{|p|} = \text{Correlation}(\text{filters}, \text{queries}_{|2|})$;
   6. $w_{|p|} = \text{Correlation}(\text{features}, \text{queries}_{|3|})$;
   7. $w_{|p|} = \text{Correlation}(\text{features}, \text{queries}_{|4|})$;
   8. normalizedPerformanceMetricsRatings(timeSpent, scrolling, click, numClicks);
   9. $r_i = (\text{w}_1 \times \text{wTime}) + (\text{w}_2 \times \text{wClick}) + (\text{w}_3 \times \text{wNumClicks})$;
   10. finalScores = $\sum_i r_i$;
11. return buildRecommendation(finalScores, k);

**Method:** buildRecommendation(finalScores, k);

1. retrieve from database k queries with the highest scores, split into terms, filters and features
2. create a string with unique terms
3. create a string of unique filters taking care about their compatibility
4. create a string with unique features
5. join terms, filters and features into one string to create a query
6. return query

---

**C. Architecture**

In this work PRISENIT is designed based on Apache Solr\(^1\). Management data files are constantly retrieved from a managed network and then indexed for searching. The data includes fault data (FM), configuration actions (CM), and performance data (PM). Data files are retrieved from the Hadoop\(^2\) RDBMSs and XML data stores used in the Ericsson system. Fig. 2 illustrates the architecture of the PRISENIT recommendation system within the overall architecture.

The main feature of this kind of system is the ability to predict the additional search actions or steps that may be required by users to extract required information. As a result, all useful additional steps that may be required by a user are suggested and can be applied in a single action. PRISENIT learns previous user search interactions and experiences to give suggestions for new search queries.
Network management data has certain characteristics that affect its usefulness: Certain network events may only be valuable within a certain time range before going stale (e.g., last 5 minutes); Alarms may be sorted by severity; Search results may be grounded to make sense; and some keyword combinations may not have matching results. In summary, the key advantages are:

- Reducing the complexity and time (interaction) needed for users to find final results and to improve search efficiency
- Reducing the use of system resources by avoiding unnecessary queries

### IV. System Application

The main goal of this search architecture is to assist the user to find and solve network problems with fewer interactions with the search interface. To demonstrate a prototype front-end search interface, we used third party web interface called HUE\(^1\) from Cloudera. We developed a new standard search interface, and a new widget to optionally view and use recommended queries recommended by the recommender system. The search system can then be invoked with either the user’s query or the recommended query.

Data that is connected with the user and the queries are stored in a RDBMS database. We also have Ericsson network management data indexed in Apache Solr to test the system and see the results and possible recommendations. We choose Solr instead of an RDBMS because existing management data is not updated (read-only), rather new data is constantly inserted. In this scenario an indexing mechanism is more effective than a traditional database approach, thus allowing faster data retrieval.

The recommender module works as an independent plugin. If it is activated, it returns and presents recommendations depending on the query that the user performed. This is a pluggable and extensible approach. After every search, the user will be presented with an optional recommendation and different results, facets, charts and so on depends on what was chosen before. If the user clicks on a chart, a filter is applied, thus creating a new query to Solr. Therefore, the user will then have an updated recommendation.

As an example scenario: In a mobile telecommunications network a Radio Network Controller node (RNC) is responsible for controlling and managing multiple base stations in a wireless network (UMTS). If a node (network element RNC75) has a problem, and assuming that previous troubleshooting the user usually finds that the problem is somehow liked to a problem in node RNC85, the system will over time learn and create an association between RNC75 and RNC85. This means that the next time that the user searches for RNC75 the system will recommend examining data from RNC75 and RNC85 to troubleshooting the problem.

Another possible scenario is when an OSS engineer is searching for alarms. If she usually sorts the results by priority, next time the system will recommend this sort criterion.

---

\(^1\)http://gethue.com/
The main advantage with this system is that the user can find the optimal result with fewer iterations. In the previous system the user needed to look for a specific RNC or fault. Then, in the result section, he needed to examine partial results and check if there were any other RNCs correlated with this issue. In addition the user would have to apply the filters that she wants. With PRISENIT, such interactions are learned and next time offered to the user. The only thing that the user has to do is to select the recommended search.

V. EXPERIMENTAL EVALUATION

For the evaluation of PRISENIT, we used data from an alarm list, performance counter values, logs and performance event logs derived from a text network in an Ericsson test lab.

A. Performance

For all evaluations we used a Ubuntu 12.04 (precise) 64-bit system with Intel® Core™ i5-2540M CPU @ 2.6GHz x 4 and 4GB of RAM.

Fig. 3 shows the time that the system needed to create a new suggestion for the user, depending on how many queries are analyzed. The experiments were conducted multiple times and averaged in order to compensate for other random factors that may influence the response times (such as usage of the system). In this case we can appreciate that the increase of response time of the system is somewhat linearly related to the number of existing queries stored in the database. According to [16], the tolerable waiting time it is around 2 seconds, so the response time of system remains acceptable. In real network management deployments, there should be little need to examine queries more than a few days old, so a limitation to examine only several thousand existing relevant queries is reasonable.

In our deployments the terms are often short and tend to be composed by one or two keywords. In this system, the filter part of the queries often has more data. For example, in our system there are many facets to show results grouped by specific fields with specific field filters, such as sorting, limits, type and so on.

B. Usability and usefulness

We conduct our usability assessment using a questionnaire. In this case, usability focuses on the effectiveness of the system for network intelligence and troubleshooting from the point of view of the end user (network engineers). We select the System Usability Scale (SUS) [17] because it is technology agnostic, making it flexible enough to assess a wide range of interface technologies, and allows quick and easy assessment of the usability of a given system.

In [18] it is stated that a sample of 12 users is sufficient to provide correct findings (i.e. the same score as a larger sample) 90–100% of the time and a sample of only 10 users, 75–80% of the time. Moreover, the original SUS instrument is composed of 10 statements that are scored on a 5-point scale of strength of agreement, among which odd questions are positive statements and the rest are negative. Final scores for the SUS can range from 0 to 100, where higher scores indicate better usability. Table 3 shows our SUS results.

<table>
<thead>
<tr>
<th>Queries</th>
<th>Users</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1</td>
<td></td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4.25</td>
</tr>
<tr>
<td>q2</td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1.67</td>
</tr>
<tr>
<td>q3</td>
<td></td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4.17</td>
</tr>
<tr>
<td>q4</td>
<td></td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2.17</td>
</tr>
<tr>
<td>q5</td>
<td></td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4.67</td>
</tr>
<tr>
<td>q6</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.08</td>
</tr>
<tr>
<td>q7</td>
<td></td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.42</td>
</tr>
<tr>
<td>q8</td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>q9</td>
<td></td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4.33</td>
</tr>
<tr>
<td>q10</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1.83</td>
</tr>
<tr>
<td><strong>Scores</strong></td>
<td><strong>80</strong></td>
<td><strong>97.5</strong></td>
<td><strong>70</strong></td>
<td><strong>75</strong></td>
<td><strong>70</strong></td>
<td><strong>92.5</strong></td>
<td><strong>90</strong></td>
<td><strong>92.5</strong></td>
<td><strong>75</strong></td>
<td><strong>90</strong></td>
<td><strong>87.5</strong></td>
<td><strong>87.5</strong></td>
<td><strong>83.96</strong></td>
<td></td>
</tr>
</tbody>
</table>
The survey was completed by network engineers. Evaluating our system, Table 3 shows an average score of 83.96. When converted to percentiles [19] this score means that we can get a grade A for our system. Using the scale used in [20] the grade of our system would be grade B. Moreover, the system would be described as “good” and would be considered acceptable. Besides, some of the participants (OSS engineers) point out that this system was much better that what they have at this moment. They said that it is easier to use, more intuitive and the recommender part can help them for troubleshooting.

VI. CONCLUSION AND FUTURE WORK

The main purpose of this work is to develop a system to improve search efficiency for network monitoring and troubleshooting and to make user experience easier and more efficient by reducing costly human interaction with the system.

While this paper focuses on finding a comprehensive and robust approach to calculate the similarity between user interactions, such an approach will contribute significantly in a wide variety of telecom network management use cases apart from alarm filtering and correlation. The system is capable of dynamically creating associations between attributes. Furthermore, the system updates these associations as the user necessities are changing.

Moreover, this system is extendable and pluggable. The actual system can work without the recommender system. Hence, the user can choose if he wants to see a recommendation or not.

From our evaluation, we saw that the overhead of the system is somewhat small with the number of previous searches stored. Therefore it will be necessary to create a mechanism to periodically remove searches that are found not to be reusable or widely relevant.

It would also be interesting to compare this approach to one where recommendations are generated for terms, filters and features separately, rather than generating an entire query. This could be achieved by creating separate recommendations with single action-based weightings for terms, filters and features. On the other hand, another possible approach would be to use different weightings to increase the importance of term similarity rather than filter/feature similarities.

It is also planned to enhance this approach with a prediction mechanism to proactively present recommended searches before or as interesting faults or events occur in the network.

REFERENCES


