COLLABORATION OPPORTUNITY FINDER

Damjan Demšar[1], Igor Mozetič[1], Nada Lavrač[1,2]

{damjan.demsar, igor.mozetic, nada.lavrac}@ijs.si
1 Jožef Stefan Institute, Ljubljana, Slovenia
2 University of Nova Gorica, Nova Gorica, Slovenia

We have designed and implemented a software tool coFinder – a collaboration opportunity finder - aimed at facilitating the work of an opportunity broker in Collaborative Networked Organizations. It crawls the web to find potential collaboration opportunities (calls for tenders) and compares them with the actual competencies of companies from a Virtual organization Breeding Environment (VBE). The coFinder tool then structures, aligns and matches textual descriptions of calls for tenders with competencies and suggests a list of potential collaboration opportunities to the broker.

1. INTRODUCTION

A Virtual organisation Breeding Environment (VBE) (Camarinha-Matos and Afsarmanesh, 2003, Plisson et al. 2007) is a long-term association of organisations and related supporting institutions with the potential and willingness to cooperate. A Virtual Organisation (VO) is a short-term alliance of organisations from a VBE with a specific goal of acquiring and fulfilling a business opportunity.

We have designed and implemented a software tool, coFinder, which helps the VO broker in identifying appropriate business opportunities from public Calls for Tenders (CfT). The coFinder tool is aimed at facilitating the work of a VO broker. It uses the same approach that is usually carried out manually by the broker: comparing potential collaboration opportunities (CfTs) with the actual competencies of the VBE, stored in the Profiling and Competency Management System (PCMS, which is part of the VBE Management System VMS) (Ermilova et al., 2005). In order to automate this process, the coFinder tool needs comparable structure of information contained in both sides. These structures can then be aligned and matched with each other in order to detect similarities and consequently detect possible collaboration opportunities. The matching in the coFinder tool is based on the comparison of textual descriptions of tenders and VBE competencies. Like the broker, the tool is able to browse public CfTs available on the web and extract tenders' descriptions from the relevant web pages. Similarly, competencies are also described in web pages or can be manually entered in text format within the coFinder tool. Eventually, coFinder will interact with the PCMS in order to get the competencies. Once the tenders' descriptions and competencies have been provided, coFinder is able to compute their similarity in order to estimate the interestingness of tenders and identify the most promising ones, and finally to propose them to the broker as potential collaboration opportunities. The output of the system is a list of

potential collaboration opportunities stored in an XML document. The structure of the document follows the directives imposed by an XML schema specified by the VMS.

In Section 2 we describe the overall architecture of coFinder. A detailed description of the collaboration opportunity identification is in Section 3. An example of using coFinder is given in Section 4. Section 5 presents the results of a recent evaluation of coFinder on scenarios with real-world collaborative networked organizations. In conclusion we show the planned integration of coFinder with other software tools for management of virtual organizations.
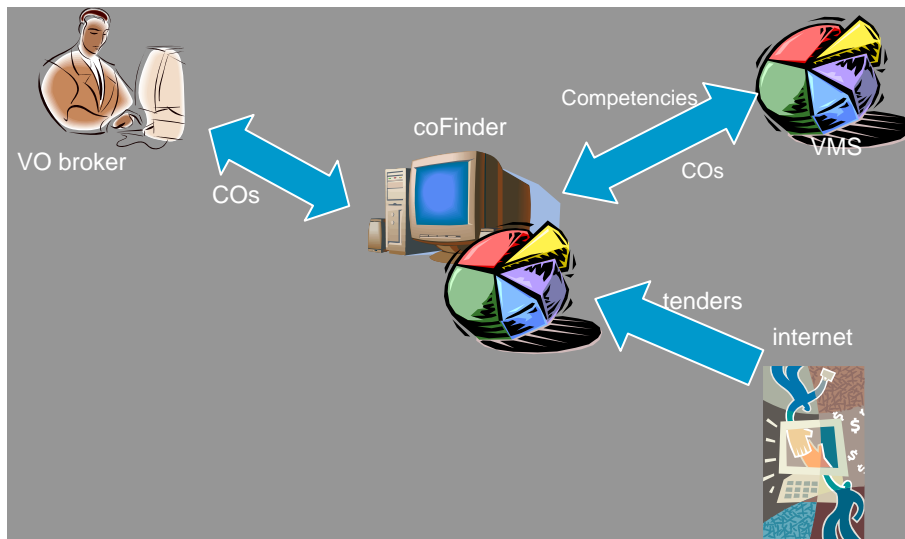
2. THE ARCHITECTURE

Figure 1: the architecture of coFinder (CO denotes a collaboration opportunity)

The architecture of coFinder is shown in Figure 1. The tool is accessed through a web interface by the VO broker. The coFinder tool accesses internet to collect the data from specified tender servers. It also accesses the VMS (PCMS) to gather competencies data, and returns the selected COs to the VMS.

Implementation-wise, coFinder is a set of PhP tools around a local MySQL database. The most important tools are crawler, parser, browser, and the CfT server template set up tool. The CfT server template set up tool is used to help the user to set up a template, which is used to separate CfTs from other web pages and collect the information from CfT web pages. Additional simple tools are used to collect other information needed to run coFinder.

Some of the tools are used only during the set up (for example the CfT server template set up tool), other tools run on a regular schedule (like crawler and parser), and the rest are used as an interface between the user and the database (for browsing and searching the data). The information always flows through the database. The initial set up data (along with the addresses of the CfT servers and templates for each of the servers) is stored in the database. It is then collected by the crawler and

used to gather all new information (web pages) from each of the CfT servers. Web pages are again stored in the database, from where they are collected by the parser. Parser uses the templates to sift out the CfT web pages from the rest and to extract the information about the CfT. This information (in an XML form), along with the calculated similarity to the VBE's competencies, is stored in the database. It is used to inform the users about new CfTs and is also available for browsing.

3. THE COLLABORATION OPPORTUNITY IDENTIFICATION PROCESS

The collaboration opportunity process comprises several steps:
1. input of necessary data, such as list of tender servers urls, templates, XML schemas, etc,
2. crawling tender servers,
3. parsing the crawled web pages,
4. matching the tender descriptions with the VBE competencies,
5. browsing, editing and adding new tender data.

Steps 1 and 5 require user interaction, while steps 2-4 are automated.

In step 1, the user has to provide some information such as a list of tender servers that he intends to use for finding potential collaboration opportunities. In addition, for each server, the user must create a template that is used to help the system identifying tenders' descriptions inside the HTML pages. Indeed, the HTML format itself does not provide any semantics to access this information directly; therefore a template is needed to detect the tender structure and its contents within the HTML page. Usually, the tender pages on the same server are generated automatically and thus share the same structure to represent tenders. Typically, one or two templates suffice to process all tenders of a single server.

There are two possible ways of creating a template:
- automatically, by comparing two existing tenders on the server, or
- manually by editing the HTML code of an existing tender.

In both cases it is possible to refine the template afterwards with one or more existing tenders or again manually by editing the code of the template.

The creation of a template takes another parameter into consideration, namely the XML schema for the output of collaboration opportunities (the CO schema). In order to create the template automatically, the system takes two existing tender pages and compares them. Since they are generated automatically, they have the same structure but different contents. Only the contents, not the form, should change from one tender to another. This makes it possible to figure out where are the variable parts of the page. A tender template is an HTML code with fields in the form of _#_fieldname_#_ representing the position of data in the tender HTML file. Field names must correspond to the field names in the CO schema. All variable pieces of the server provided HTML must be marked as a field, and as much as possible of unchangeable HTML should remain in the tender template. The template is stored along with the server data in a local database.

Another input needed from the user is the competencies of the VBE. Competencies, in the current system, are entered manually by the broker. It is planned to extend the system in the next version so that it can automatically collect

competencies from the PCMS. Competencies are organized in two categories: general competencies and specific competencies. When the connection with PCMS is established, competencies will be collected periodically in order to keep track of the changes, for instance if some partners leave or join the VBE. Eventually, additional information, such as synonyms (e.g., found in the Wordnet lexical ontology), will complete existing competencies.

In step 2, after all the above information is provided, the coFinder tool crawls all the servers specified by the user and gathers tender pages. It is possible to schedule this process to retrieve new tenders periodically. The implementation is dependent on the operating system (for instance Scheduled Tasks on Windows, or crontab on Linux).

The next, step 3, is to parse tender pages. The template built previously is used here to extract tenders' descriptions from their web page. The information is then stored in the database using the CO XML schema.

Step 4: Once the tenders' description have been extracted and structured, it is possible to identify potential collaboration opportunities by matching tenders' description with the competencies available. The interestingness factor is computed for each tender. This measure is a weighted sum of the similarity between textual fields describing the tender (description field in XML and the whole XML) and textual fields describing the VBE competencies (e.g., general and specific competencies).

Step 5: To limit the number of potential collaboration opportunities, a threshold is used for the interestingness measure. This threshold is set by the user at the beginning of the process. The VO broker will be notified by an e-mail only about the potential collaboration opportunities that are above this threshold.

It is also possible to browse and search through all the collected tender data, select the tenders suitable for VO creation, or just select the data to be send by an e-mail to the VO broker. Since all needed data is not always available in the tender data supplied by the server, the system offers the possibility to include some data that the VO broker may collect using other methods (phone calls, email correspondence, etc.). Such data can be useful in further steps of VO creation.

The coFinder structure and data flow is shown in Figure 2. CoFinder first crawls each website from the list of servers and stores the web pages in the local database. Then it parses the collected web pages using tender templates created by the user. For the resulting tender data the interestingness measure is calculated (again storing the complete result in XML form in the local database), and finally the VO broker is notified. The VO broker then checks and possibly edits the data and selects the tenders which are suitable as a basis for potential VO creation. The selected tenders (which can now be called collaboration opportunities) are then delivered as the output of coFinder.

4. AN EXAMPLE OF USE

In this section we show an example of the use of coFinder. When the tool is used for the first time, initial data has to be set up. Most of that data is a simple text like VBE competencies, the e-mail addresses for notification, and the URL addresses of CfTs servers. The most important part of the setup is the definition of CfT server templates. The template is usually generated from two or more CfTs on the server, using a tool in coFinder. This tool compares the web pages of CfTs and

asks the user to define the meaning of the differences. The tool in use can be seen in Figure 3, a CfT provided as an input to the tool in Figure 4, and the final template in Figure 5.
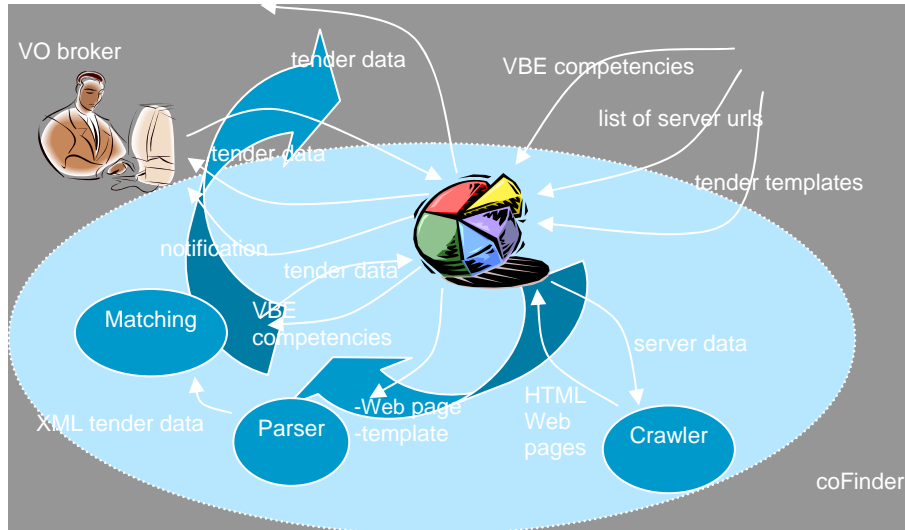

Figure 2: coFinder working procedure

Figure 3: An example of template generation tool in coFinder

After the initial data has been set up, coFinder crawls the specified servers, parses the collected web pages and puts the information about the collected CfTs into the internal database. All CfTs are compared to the VBE competencies as they were set up at the beginning and interestingness factor is calculated. A notification by e-mail is then sent to the specified address, giving information about the best CfTs found on the servers (if any is above the pre-specified interestingness threshold).

The user/broker can use coFinder to browse and search through the collected CfTs (Figure 6). Both, the complete data of the selected call for tender (Figure 7), as well as the original web page can be inspected. The data can also be edited in order to add additional information collected by the broker (e.g., through e-mail, phone calls or other methods) and thereby enrich the data needed during further steps of the VO creation process. To start the process, the broker selects the appropriate CfTs, and marks them as appropriate for VO creation. coFinder then makes the data about appropriate CfTs available to subsequent tools that plan the process needed in VO, suggest the partners, help with negotiation…

Figure 4: An example of Call for Tender from the TerndersInfo.com server

Figure 5: The generated template for the TendersInfo.com server

5.   EVALUATION

The coFinder tool was evaluated by three groups of end-users (all from existing VBEs: Swiss Microtech, CeBe Network and ISOIN) after a hands-on exercise. Most of the end-users agree that coFinder is useful in real-life business and will help to save time in their operations. Table 1 shows a part of evaluation questionnaire with

answers of one of the end-users. However, most end-users also agree that coFinder should be made easier to use, so we still have to improve on its user interface.

Figure 6: Searching through collected CfTs

Table 1. Part of the end-user questionnaire with answers of one of the test users.

Question	Response
Functionality tested :	Co-Finder
The material objectives are easy to understand	Agree
What makes the comprehension of the material particularly easy ?	CFT are collected automatically
What could be done to increase the comprehension easiness of this material ?	Attach Reference Guide
The material is really useful in your real-life business	Strongly agree
It will help you to gain time in your business	Strongly agree
It will professionalize your activity	Strongly agree
It will increase the trust in your collaborative organization	Agree
It will decrease the risks in your collaborative organization	Disagree
The use of this material will change your business management	Agree
Which are the three parts/functionalities of the material you find particularly relevant ?	Automation of search / Manually create CFT/ Report Status of each CFT
The material is easy to use	Agree
The vocabulary used in the material is relevant and clear	Disagree
The described methodology can be used in your organization	Agree
This template will be applicable in your business environment	Strongly agree
Functionalities are easy to find	Agree
Functionalities are easy to understand	Disagree
Functionalities are easy to handle	Disagree

```
CO_customer_name        ation Details Organisation CHINA INTERNATIONAL TENDERING COMPANY
CO_customer_URL
CO_customer_address     CHINA INTERNATIONAL TENDERING COMPANY ROOM 514, JIU LING BLDG. (NORTH WING), NO. 21 XI SAN HUAN BEI LU
                        BEIJING 100089 CHINA Tel : 86-10-68404404 Fax : 86-10-68316696
CO_customer_country     UNKNOWN
CO_CP_name
CO_CP_phone
CO_CP_fax
CO_CP_email
CO_title                UNKNOWN
CO_language
CO_description          SUPPLY OF THE FOLLOWING EQUIPMENTS:- PROFESSIONAL GIS; DESKTOP GIS; GPS; AUTOMATIC WEATHER STATION;
                        AUTOMATIC RAIN-GAUGE; AUTOMATIC RAIN-GAUGE; AUTOMATIC WATER LEVEL; REMOTE SENSING IMAGE; REMOTE SENSING
                        IMAGE; SOFTWARE FOR IMAGE PROCESSING; AUTOMATIC WATER SAMPLER; TDR SOIL MOISTURE METER; ELECTRONIC
                        BALANCE; DRYING OVEN; RAINFALL SIMULATOR; TRANSPORT VEHICLE; COMPUTER; NET SERVER; SWITCH; NOTEBOOK
                        COMPUTER; SCANNER; COPIER; PAGE SETTER; BINDING MACHINE; PRINTER; SLIDE PROJECTOR; PROJECTOR ETC.
CO_project_name         BEIJING ENVIRONMENTAL IMPROVEMENT PROJECT [(LOAN NO.1336) LOAN FUNDED BY THE ASIAN DEVELOPMENT BANK (ADB)]
CO_sector               UNKNOWN
CO_document_type        wbg"> Tender Notice
CO_bidding_type         International Competitive Bidding
CO_contract_type
CO_target_price         UNKNOWN
CO_objectives           UNKNOWN
CO_duration             UNKNOWN
CO_issuing_date
CO_opening_date
CO_due_date             Not Provided
CO_closing_date
CO_closing_time
CO_product_description
CO_volume               UNKNOWN
```

Figure 7: coFinder parsed data from the call for tender shown in Figure 4

6.   CONCLUSION

The coFinder tool is a software tool which can support a Virtual Organization broker in the management of the Virtual Breeding Environment. Its main goal is to help the broker in finding appropriate collaboration opportunities which match the competencies of the VBE. The selected potential collaborative opportunity can be passed to other CO management software tools, such as a tool for collaborative opportunity characterization and rough planning, followed by a tool for appropriate partner search and selection, which can be invoked to suggest a list of potential companies from VBE to actually fulfill the selected collaborative opportunity. In parallel to these activities, there is a negotiation wizard to assist during the negotiation processes of the VO contract.

Acknowledgments

7.   REFERENCES

1. Camarinha-Matos LM, Afsarmanesh H. Elements of a base VE infrastructure. Computers in Industry, vol. 51, pp. 139-163, 2003.
2. Plisson J, Ljubič P, Mozetič I, Lavrač N. An ontology for Virtual Organization Breeding Environments. To appear in IEEE Trans. on Systems, Man, and Cybernetics, 2007.
3. Ermilova E., Galeano N., Afsarmanesh H.,ECOLEAD deliverable D21.2a Specification of the VBE competency/profile management, 2005.