

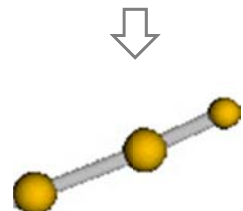
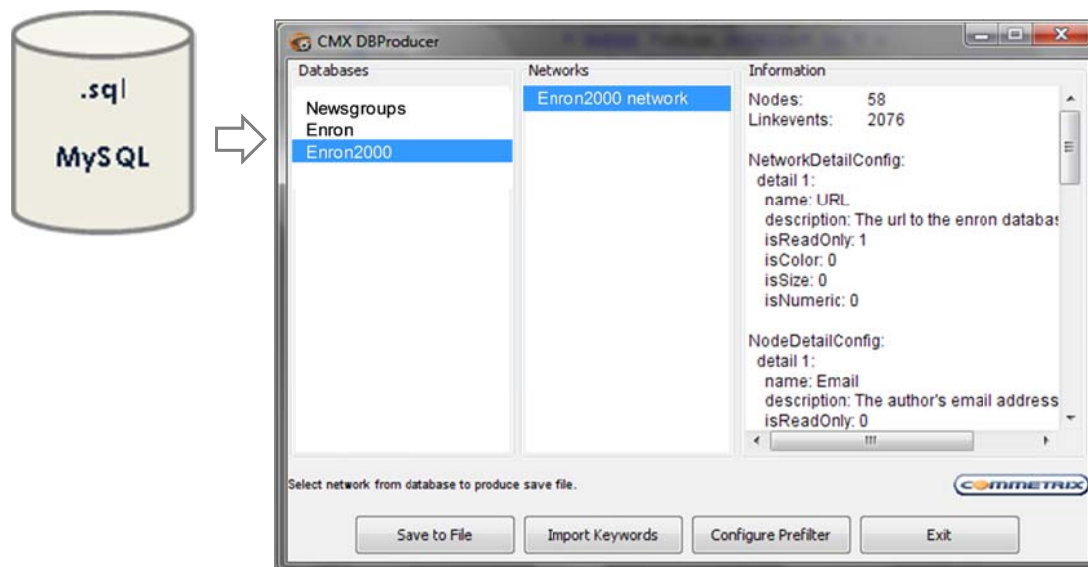


USER MANUAL
VERSION 2012-05-02



Commetrix DB Producer Producing CMX Files from MySQL Databases

Product Website: www.commetrix.net
Vendor Contact : info@trilexis.com



.cmx for Commetrix Analyzer

Contents

1	Introduction	3
2	Licensing Commetrix DB Producer.....	3
3	Installation of Commetrix DB Producer	3
3.1	Java	4
3.2	MySQL Community Server.....	4
3.3	MySQL J/Connector.....	4
4	Logging into the Commetrix DB Producer	4
5	Importing from a Commetrix MySQL Database	5
6	Configuration of PreFilters.....	6
6.1	Isolates PreFilter	7
6.2	Min Linkevents PreFilter	7
6.3	Min Keyword Length PreFilter	7
6.4	Linkevent Detail as Keyword PreFilter	8
6.5	Node Detail as Keyword PreFilter	8
6.6	Min and Max Time PreFilter	8
6.7	Keyword PreFilter	8
6.8	Node Replacement PreFilter	8
7	Other Import Opportunities	9
8	Contact	9

1 Introduction

COMMETRIX is a software framework for the analysis of social networks. Next to conventional static network analysis of cumulative snapshots of relationship networks, its event-based network model enables the examination of network evolution over time as well as the study of link and node properties.

Commetrix offers an innovative **DATA MODEL**. Unlike most other SNA tools, it does not store links directly as valued relationships but as individual relational events. This provides the opportunity of modelling networking processes with multiple types of relationships and comprehensive qualitative and quantitative node and link attributes in a single dataset. With this approach, users can encode aspects like topic descriptors (e.g. content coding, keywords), types of links, e.g. socialization, document exchange, affiliation, media, time stamps of links, links connecting multiple nodes, as well as arbitrary quantitative or qualitative variables classifying nodes, e.g. affiliation, age, types, etc.

The **USER INTERFACE** is very easy to use. It is developed to visually support exploratory examination of a network dataset in order to identify and observe relevant substructures, periods, and processes of your network data. Commetrix computes time window measures and additionally provides very sophisticated functionality for displaying and animating the community evolution as an evolving graph – allowing the user to visually inspect the actors' activities. The animated graph, called **communigraph**, is one of the best existing visualizations of network change. Visual variables can be set by the user to represent node and link properties by label, node size, node color (brightness, transparency), or a number of rings around the node.

Complex options for time, actor, relationship, and topic filtering help to **FOCUS RELEVANT STRUCTURES**, i.e. relevant nodes, relationships, time periods, or even topics. For any time period and for any selection, typical social network measures can be computed and analyzed. Selections can be exported to tables for further analysis. This includes the export of network changes over time. All these features help to actually represent and visually trace change in a network and adds additional insight to the quantitative results.

This manual gives an introduction to the Commetrix Producer which retrieves network data from Excel and other file formats and creates Commetrix network files (.cmx) that can be visualized by the Commetrix Analyzer.

Also note the online video introduction to producing network files with Commetrix Producer:

<http://www.youtube.com/watch?v=CxpUHrf6DNc>

2 Licensing Commetrix DB Producer

Before you acquire a license of Commetrix DB Producer (also referred to as CMXDBProducer) you can try it with smaller datasets for free (see section 4) via the [trial version](#).

If you want the [full license](#), please have a look at: <http://www.commetrix.net/lic> for documents explaining possible Commetrix-based solutions and for information about the ordering process. You can also press the order button in the login screen.

3 Installation of Commetrix DB Producer

Installation is simply done by [unpacking the downloaded .zip file into a folder of your choice](#) (we would advise not to use the location C:/Program Files, as special protections may later prevent you from storing files or changing the folder's contents). We do not guarantee the CMXProducer working properly with other operating systems than Windows XP, Windows Vista and Windows 7.

After the installation you can either use the trial mode (see section 4) or acquire a license (see section 2). In the following, further software requirements (i.e. Java) for running the CMXProducer are explained.

3.1 Java

To run the CMXDBProducer (as well as the CMXAnalyzer) a Java Virtual Machine is required. On most machines it is already installed. If not, download the Java Virtual Machine (<http://www.java.com/de/download/>) and follow the installation instructions given there. CMXProducer requires at least Java 1.6 (Java Version 6). We do not guarantee the CMXDBProducer working properly with older versions. Keep the path to the Java Runtime Environment (JRE) folder (<path-to-jre>) as it may be required later.

3.2 MySQL Community Server

To store Commetrix network data and to create CMX files from them a MySQL database is required.

Download the MySQL Community Server (<http://dev.mysql.com/downloads/>) and follow the installation instructions given there. CMXProducer requires at least MySQL 5.0. We do not guarantee the CMXProducer working properly with older versions. Keep the adminname and password combination for the MySQL database as it is required when running the CMXProducer.

To monitor the network data in the database some GUI tools like the MySQL Query Browser might be helpful. They can be downloaded at <http://dev.mysql.com/downloads/gui-tools/5.0.html>.

3.3 MySQL J/Connector

To allow a Java-based program like the CMXProducer to access a MySQL database a JDBC driver has to be installed. Download the MySQL Connector/J (<http://dev.mysql.com/downloads/connector/j/>) and follow the installation instructions given there. The connector will be a .jar file which has to be copied to <path-to-jre>/lib/ext folder of your Java Virtual Machine. CMXProducer requires J/Connector 3.0 to ensure all database functionalities. We do not guarantee the CMXProducer working properly with other versions.

4 Logging into the Commetrix DB Producer

To run Commetrix Producer, either a valid license file together with username and password is required , or the trial mode can be selected.

The trial mode requires no licensing/activation process and there is no limited trial period. Rather, in the trial mode Commetrix Producer considers all nodes of the provided dataset but only the first 50 link events (e.g. messages) among them. That means that smaller datasets can be produced completely while larger files will yield an incomplete network datafile (.cmx), which can only be used to test functionality of Commetrix Producer or Commetrix Analyzer but not for a reliable analysis.

If you decide to license the Commetrix framework, then have a look at section 2 for further guidance. After the licensing process is carried out you receive a license activation file (License_Producer.dat), a login and a password (which you can propose).

After the username and password to run the Commetrix Producer have been entered (see figure below) and checked, the main interface is presented. The next section discusses its main functions.



Figure 1: Commetrix Producer – Login Window. The Trial mode is activated with the “Trial” button. The “Order” Button leads you to web-based information about solutions, prices, and about the ordering processes.

After the username and password to run the CMXProducer have been validated, the user has to enter the database parameters: admin name and password of the MySQL database, server ip, and database port. Running the MySQL database on the same computer as the CMXProducer the normal parameter values (except the password) are shown in the figure below:



Figure 2: Commetrix Producer – Database parameters.

These configurations are saved to the `cmxproducerconfig.prop` file. If the parameters are valid the Commetrix DBProducer main GUI is shown.

5 Importing from a Commetrix MySQL Database

The Commetrix DBProducer main gui allows to choose a network out of a Commetrix database. There can be several Commetrix databases (shown in panel ‘Databases’) and each database can contain more than one network (shown in panel ‘Networks’). For each selected network some information about the number of nodes and linkevents and which details are in use are presented (panel ‘Information’).

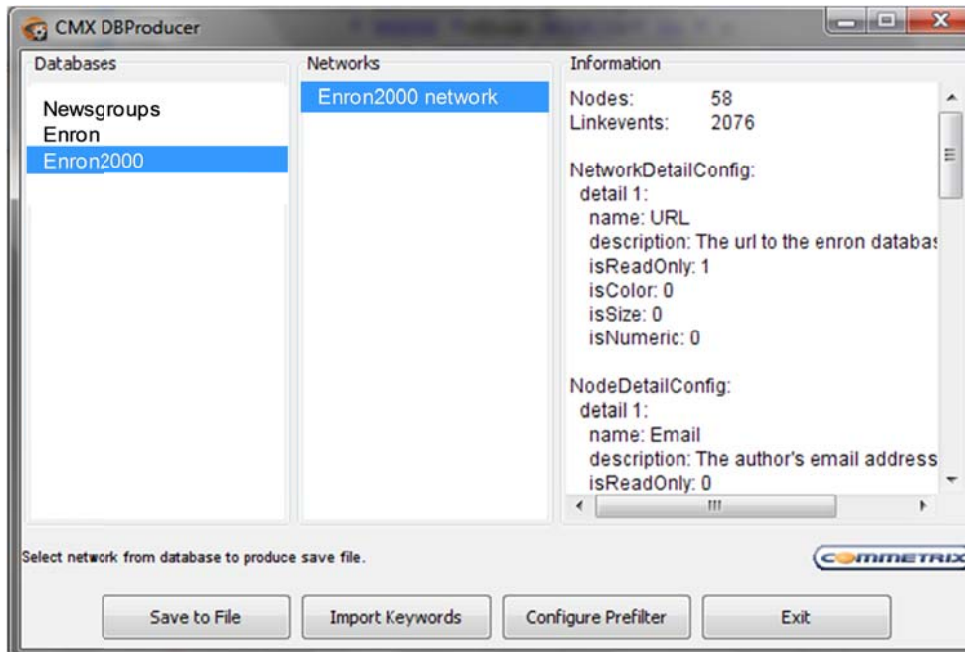


Figure 3: Commetrix Producer – Network selection GUI.

After a network is selected a Commetrix network file can be created (button ‘Save to File’). The time to process the data and create the network depends on the size of the data set. After the network is successfully created and saved to file a message is shown (above the button ‘Save to File’) where the file can be found. In general, it is saved to the `<path-to-cmxproducer.exe>/save/<database name>/<network name>_<timestamp>.cmx`.

Before the network is created the user might want to employ a text mining process. A rough text mining process is available via button ‘Import Keywords’. This process splits the entries in columns ‘Subject’ and ‘Content’ of table ‘Linkevent’ into single tokens (keywords) and removes stopwords as well as too short keywords from the data set. All other keywords are stored in table ‘Keyword’, their occurrences are stored in tables ‘SubjectKeyword’ and ‘ContentKeyword’. The stopwords retrieved from table ‘Stopword’, the minimum length of a keyword is retrieved from the Prefilter configuration file used together with the network (see next section).

The CMXProducer allows the user to employ some prefilter on the database entries. This function is available via button ‘Configuration of PreFilters’ and explained in the next section.

6 Configuration of PreFilters

The CMXProducer allows to set several prefilter, e.g. to use the data stored in a linkevent detail as keyword or to ignore isolated nodes when creating a Commetrix network file. Whenever a file is created different filter options can be set. Thus, each network stored in a Commetrix database can be analyzed from different perspectives. The design and usage of the prefilter is explained in more detail in the CMXProducer manual. During the file creation process the tables ‘Node_tmp’, ‘Linkevent_tmp’, ‘LinkeventSender_tmp’, ‘LinkeventRecipient_tmp’, ‘LinkeventParent_tmp’, ‘Keyword_tmp’, ‘SubjectKeyword_tmp’ and ‘ContentKeyword_tmp’ are created and contain only the data used for creating the Commetrix network file. They are kept in the database until the CMXProducer is used again, to allow the user to examine the data if necessary. The structure of these tables correspond to the original tables ‘Node’, ‘Linkevent’, ‘LinkeventSender’, ‘LinkeventRecipient’, ‘LinkeventParent’, ‘Keyword’,

‘SubjectKeyword’ and ‘ContentKeyword’. These tables will be created by the CMXProducer, they are not included in the database scripts for creating an empty or exemplary Commetrix database.

At the moment the following prefilter are available: isolates filter, min linkevents filter, min keyword length, linkevent detail as keyword filter, node detail as keyword filter, min or max time filter, keyword prefilter, node replacement filter.

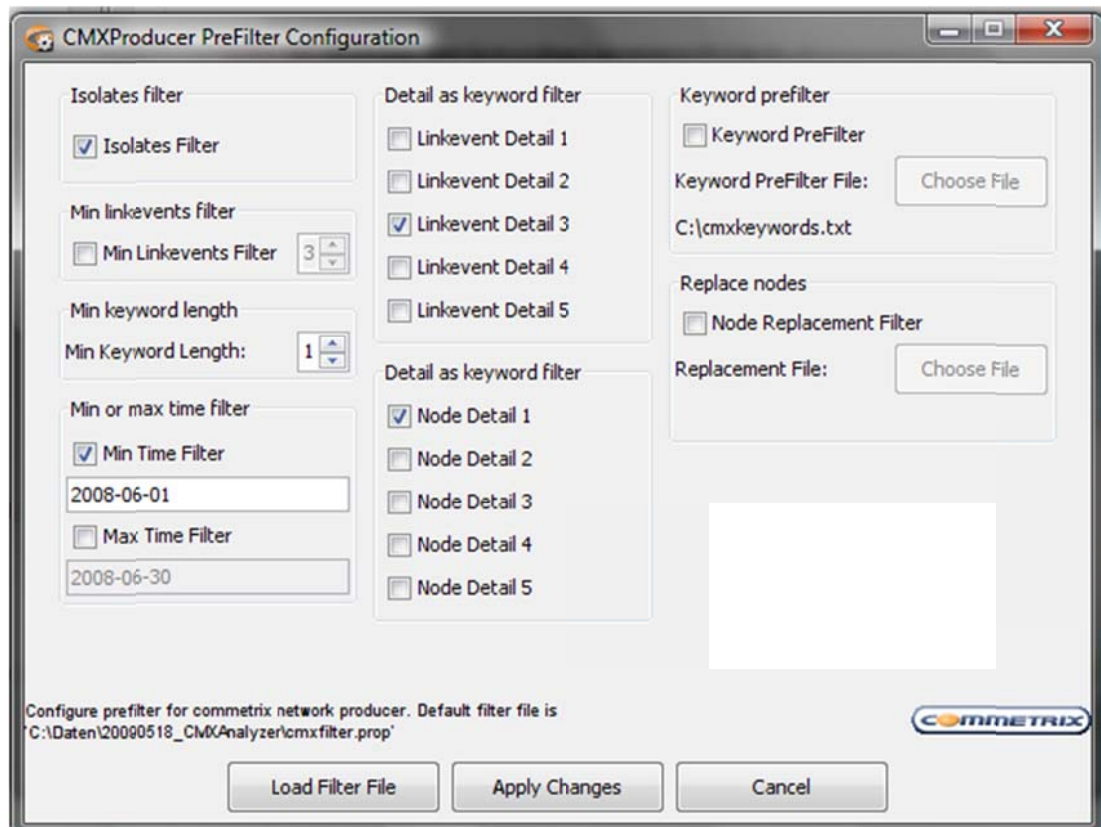


Figure 4: Commetrix Producer – PreFilter Configuration GUI.

6.1 Isolates PreFilter

If the isolates filter is chosen, only those nodes are retrieved from the database which have at least one link to another node. The isolated nodes are not removed from the database.

6.2 Min Linkevents PreFilter

If the min linkevents filter is chosen, only those nodes are retrieved from the database which have at least the minimum number of linkevents send or received which is selected.

6.3 Min Keyword Length PreFilter

If the min keyword length filter is chosen, the text mining process available via the CMXProducer main gui (button ‘Import Keywords’) removes all keywords with a keyword length less than the selected minimum length.

6.4 Linkevent Detail as Keyword PreFilter

If a linkevent detail as keyword filter is selected the data stored in the corresponding column in table 'Linkevent' is used as keyword as well. To distinguish these artificial keyword from the natural keywords of the dataset the keyword is written with a leading identifier '!LX!' where X indicates the detail used, e.g. '!L1!', '!L2!', etc. These artificial keywords are included in the table 'Keyword_tmp'. Their occurrences are set to 1.0 in each linkevent and stored in table 'ContentKeyword_tmp'. Employing the linkevent as detail filter the CMXAnalyzer will list these artificial keyword in the KeywordFilter as well.

6.5 Node Detail as Keyword PreFilter

If a node detail as keyword filter is selected the data stored in the corresponding column in table 'Node' is used as keyword as well. To distinguish these artificial keyword from the natural keywords of the dataset the keyword is written with a leading identifier '!NX!' where X indicates the detail used, e.g. '!N1!', '!N2!', etc. These artificial linkevents are included in the table 'Keyword_tmp'. Their occurrences are set to 1.0 in each linkevent assigned to the node (either having sent or received the linkevent) and stored in table 'ContentKeyword_tmp'. Employing the node as detail filter the CMXAnalyzer will list these artificial keyword in the KeywordFilter as well.

6.6 Min and Max Time PreFilter

If the min time filter is selected, all linkevents before the specified date are ignored while creating the network. The input format has to be YYYY-MM-DD.

If the max time filter is selected, all linkevents after the specified date are ignored while creating the network. The input format has to be YYYY-MM-DD.

6.7 Keyword PreFilter

If the keyword prefilter is selected, only those linkevents containing keywords which are included in the specified file are used to create the network. The file has to be a .txt file where each row contains a keyword. At the moment, the keyword prefilter cannot be used for linkevent or node details as keyword (artificial keywords).

6.8 Node Replacement PreFilter

If the node replacement filter is selected, nodes in the database can be (temporarily) replaced by other nodes or completely, e.g. if there are several nodes which should be aggregated to a single node. All corresponding linkevents will be assigned to the replacing node. The file has to be a .csv file with three columns (see table below): the first column is database id ('NodeID') of the node which will be replaced, the second column is the database id ('ReplacedbyNodeID') of the node which is used to replace the first node, the third column ('Delete_if_1') indicates whether the replaced node should be deleted from the scope of the producer. The nodes to be replaced will not be removed from the original database tables but only from the current scope of the producer. The first row (header) in the file will be ignored by the producer.

Table 1: Node replacement file – Example.

NodeID	ReplacebyNodeID	Delete_if_1
6901	764	1
5559	962	1
1312	974	1
2212	974	1
1065	1052	1
1326	1056	1
1868	1056	1
2025	1109	1
2211	1109	1
1051	1291	1
1972	1299	1
2157	1907	1
7269	2398	1
1929		1
2039		1
3213		1
4448		1
6957		1

7 Other Import Opportunities

Commetrix network datafiles can also be imported and produced from Excel tables (.xls), Comma Separated Value (.csv) Files (Tables), or from a special Commetrix XML Format (.xml). However, for this import you need the software Commetrix Producer. Please have a look at the Commetrix Producer manual for more information.

8 Contact

If our Commetrix Framework Solution is of interest for you, we will be happy to arrange a license contract. Just contact us via: info@trilexis.com (preferred) or via the following address.



Trilexis Information Management GmbH
 Helmut-Kaeutner-Weg 5,
 13591 Berlin
 GERMANY
 Corp. Registration ID: DE 29 039 00437
info@trilexis.com

