

IBM Industry Customer Intelligence Solutions
Version 1.0.0

*Behavior Based Customer Insight for
Communication Service Providers
Solution Guide*

IBM

Note

Before using this information and the product it supports, read the information in "Notices" on page 45.

Product Information

This document applies to IBM Industry Customer Intelligence Solutions Version 1.0.0 and may also apply to subsequent releases.

Licensed Materials - Property of IBM

© **Copyright IBM Corporation 2015.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Introduction	v
Chapter 1. Overview	1
Chapter 2. Installation and configuration.	5
Prerequisites	5
Contents of the solution package	5
Overview of the installation and configuration process	6
Installing the solution installer	6
Deploying the installation files on the nodes.	7
Installing the TELCO database	8
Deploying content on the Business Intelligence node	8
Creating a data source connection to the TELCO database	9
Setting up the Cognos Framework Manager model	9
Deploying the IBM Cognos BI reports	10
Deploying IBM SPSS content	11
SPSS job sequencing and triggering	12
Changing churn factor labels	12
Improving the SPSS model performance.	13
The Rapid Analytics platform	14
Installing the framework layer	14
Configuring the Rapid Analytics node	15
Configuring table groups	16
Starting the Rapid Analytics platform server	18
Installing the SLM tagging functionality	18
Installation and configuration of the Mediation node	19
Installing the mediation layer	19
Installing the mediation components	20
Configure the vedia mediation components.	21
Configuring IMSI segments on the Mediation node	21
Configuring the Gn and LTE probe collectors	22
Configuring the rapid voice and SMS probe	23
Configuring the probe traffic type	24
Configuring a roamers filter	24
Configuring a cells filter	25
Configuring the Rapid Analytics platform for TDR processing	25
Loading data and starting the predictive analytics job	26
Chapter 3. Using Behavior Based Customer Insight for Communication Service Providers	29
Using the NPS correlation application	29
Using the churn management application	31
Appendix A. Accessibility features.	35
Appendix B. The data provision tool	37
Appendix C. Troubleshooting	39
Troubleshooting a problem	39
Troubleshooting resources	39
Troubleshooting known data provision tool error messages	41
Troubleshooting the FTP collection of files in Rapid Analytics	42
Troubleshooting SPSS	44
After running the solution installer, Chef processes are still running after executing the cleanup.sh script	44

Notices	45
Index	49

Introduction

The IBM Behavior Based Customer Insight for Communication Service Providers solution gives you the information and insight that you need to provide proactive service to your customers.

Audience

This guide is intended for administrators and users of the Behavior Based Customer Insight for Communication Service Providers solution. It provides information on installation, configuration, and use of the Behavior Based Customer Insight for Communication Service Providers solution.

Finding information

To find product documentation on the web, including all translated documentation, access IBM® Knowledge Center (<http://www.ibm.com/support/knowledgecenter>).

Accessibility features

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products. Some of the components included in the Behavior Based Customer Insight for Communication Service Providers have accessibility features. For more information, see Appendix A, “Accessibility features,” on page 35.

Behavior Based Customer Insight for Communication Service Providers HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.

Forward-looking statements

This documentation describes the current functionality of the product. References to items that are not currently available may be included. No implication of any future availability should be inferred. Any such references are not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of features or functionality remain at the sole discretion of IBM.

Samples disclaimer

Sample files may contain fictional data manually or machine generated, factual data compiled from academic or public sources, or data used with permission of the copyright holder, for use as sample data to develop sample applications. Product names referenced may be the trademarks of their respective owners. Unauthorized duplication is prohibited.

Chapter 1. Overview

The IBM Behavior Based Customer Insight for Communication Service Providers is a prepackaged solution that works with IBM Predictive Customer Intelligence.

Important: Before you install the application, IBM Predictive Customer Intelligence must be installed on your system.

IBM Behavior Based Customer Insight for Communication Service Providers is a solution that drives improved marketing and customer care performance through advanced and predictive analytics of customer activity across locations, devices, applications, and interests. It highlights behavior-based information about customers using deep analysis of large volumes of raw mobile usage data, including internet usage and voice, and SMS activity.

The solution also provides a churn management application, utilizing standard churn modelling factors, as well as voice, SMS, and mobile data records, to analyze trends in subscriber voice, SMS, and mobile data usage and to interpret the churn model.

What need does the solution address?

Customer experience teams have many unanswered questions due to limited data and analytics. Some of these questions include:

- Why are my customers not satisfied or happy?
- Why are our NPS® (or T-NPS) so low?
- What segment of customers is impacting NPS ?
- What are the specific issues potentially impacting NPS and customer satisfaction?
- What can be done to improve customers' satisfaction and happiness?
- What are the factors impacting subscriber churn?
- Who are my most valuable customers?

To help address these questions, IBM Behavior Based Customer Insight for Communication Service Providers is composed of three separate components:

- an NPS correlation application
- a churn management application
- a customer behavior dataset

What does the NPS correlation application provide?

The solution helps customer experience team members in the telecommunications industry to determine customer insights by drilling down into Net Promoter Score (NPS) data and providing filters on key factors. The NPS index measures the likeliness of customers to recommend a company's offerings. It is used to gauge a customer's overall satisfaction and brand loyalty.

IBM Behavior Based Customer Insight for Communication Service Providers analyzes subscribers' relative satisfaction or happiness with their service based on

NPS. The solution allows users to drill down into the data to determine specific issues that correlate with negative NPS . The solution's metrics can be used to:

- proactively address key issues, impacting churn and other issues
- measure and correlate NPS with subscriber experience
- improve customer experience and reduce attrition by improving their NPS
- provide faster resolution of customer issues to improve customer satisfaction and loyalty

The NPS correlation application includes an IBM Cognos user interface, which is described in the workflow procedure:

- A dashboard of prioritized factors impacting transactional-NPS (T-NPS)
- The ability to quickly view the top factors that negatively affect T-NPS scores within appropriately grouped categories, such as customer, data activity, device, and location.
- Access to various data sources to build the profiles of a typical promoter and detractor, which can be used to infer the profiles' impact across the customer base, gauging NPS for all customers based on similar experience.
- Combining of data sets such as 2G, 3G, 4G LTE data, NPS survey results, cell, CRM, and device provisioning to create a single Subscriber Analytics Record for each subscriber.
- Predictive models, which correlate subscriber analytics records to identify and address possible causes for poor NPS.

The solution is delivered with pre-defined industry data models, use case specific metrics, and interactive dashboards.

What does the churn management application provide?

The churn management application provides an analysis of subscriber data suggesting how likely and when a customer is likely to churn. The management application includes an IBM Cognos user interface, which is described in the workflow procedure:

- A churn overview providing a relative segmentation of subscribers, based on their churn likelihood.
- The ability to filter subscribers based on key factors, including churn propensity, customer value (a "keep score"), predicted churn date, churn factors, and contract type.

The churn management application receives inputs from the IBM Customer Experience Analytics for Communication Service Providers solution, which helps customer experience team members to understand the subscriber experience trends across different "experience" categories, such as mobile data activity (including connectivity and retransmissions), voice, SMS, web browsing, and video. These inputs provide important insights into a subscriber's experience and are valuable when attempting to predict churn. Furthermore, not all subscribers are equally valuable to Communication Service Providers. Therefore, this solution provides a "keep score" for each subscriber, using a scale of 1-100, with 100 being the most valuable.

What does the customer behavior dataset provide?

The customer behavior dataset is a set of tables customers can access, along with an accompanying data dictionary. The dataset contains subscribers' usage data, including their interests and preferences. The application analyzes the amount of

time subscribers spend browsing, searching, reading, and streaming information on their devices. Each period of activity is categorized, where possible, into one of hundreds of interest categories, enabling communication service providers to categorize their subscribers into hundreds of different micro-segments based upon these interests.

Chapter 2. Installation and configuration

The IBM Behavior Based Customer Insight for Communication Service Providers solution is for use with IBM Predictive Customer Intelligence.

Behavior Based Customer Insight for Communication Service Providers consists of the following nodes:

- A Rapid Analytics node.
- A Mediation node.
- A Business Intelligence node.
- A Predictive Analytics node.
- A data node.

Note: The IBM Predictive Customer Intelligence topology includes an Integration Bus node, which is not used in the Behavior Based Customer Insight for Communication Service Providers solution.

Prerequisites

Before you install the Behavior Based Customer Insight for Communication Service Providers solution, there are prerequisites that must be met.

Ensure that the following conditions are met before you install the solution:

- IBM Predictive Customer Intelligence is installed and configured and you have applied all known fix packs to it.

Important: If you are using Cognos Version 10.2.1, you must ensure that you have applied the latest fix pack.

- IBM InfoSphere® BigInsights® 3.0.2, or later, is installed and configured on the Rapid Analytics node.
- You have a Linux system that you can install to. The Solution Installer works only on Linux.
- You understand your existing environment so that you can determine where you want the various components of the solution to reside.
- You have administration rights and the ability to copy files between computers.

In the context of the Behavior Based Customer Insight for Communication Service Providers solution, the BigInsights master node is the Rapid Analytics node.

Contents of the solution package

The IBM Behavior Based Customer Insight for Communication Service Providers solution package contains several parts.

The following parts are included in the solution package:

- Solution Installer
- Rapid Analytics platform
- IBM SPSS® Collaboration and Deployment Services

- IBM Cognos® Business Intelligence (BI) reports and IBM Cognos Framework Manager models

Overview of the installation and configuration process

This is an overview of the general tasks that you must perform to install and configure the Behavior Based Customer Insight for Communication Service Providers solution.

Perform the following tasks:

1. Download the solution from IBM Passport Advantage®.
2. Plan where you want the various parts of the solution to reside. For example, you must decide which computers to use for the Rapid Analytics node and the Mediation node.
3. Run the Solution Installer to deploy the installation files to the nodes.
4. Install the TELCO database.
5. Set up the IBM Cognos Framework Manager model.
6. Install IBM Cognos reports.
7. Install and configure the Rapid Analytics node.
8. Install and configure the Mediation node.
9. Configure the Behavior Based Customer Insight for Communication Service Providers database.

Installing the solution installer

Before you can configure and deploy the nodes in your solution, you have to install the Solution Installer.

About this task

Here's a short video that shows you how to install the Solution Installer:

<https://youtu.be/9u6ZOk8iyt4>

Procedure

1. Log on to the computer where you downloaded the Behavior Based Customer Insight for Communication Service Providers solution package.
2. Decompress the solution package.
3. Go to the solution installer directory in the following location:
opt/IBM/IS_CSP_Customerinsight_1.0/SolutionInstaller.
4. If you have a firewall enabled, enter the following command:
`./firewall.sh`

The ports that are used by the Solution Installer to transfer files to the target nodes are opened on the server.

5. Enter the following command:
`./setup.sh username first name last name email`

The web server is started and the URL for the Solution Installer displays in the command line window.

6. Copy and paste the URL into a web browser and bookmark it.

What to do next

Use the Solution Installer to deploy the installation files to the nodes in your solution.

Deploying the installation files on the nodes

You use the Solution Installer to deploy the installation files to the appropriate nodes in your solution.

Before you begin

Ensure that you know the URL for accessing the Solution Installer.

Ensure that you know the following information for each of the nodes in your solution:

- Host name or IP address
- A user ID and password that has write permissions to the /opt/ directory.

Each node must also meet the following requirements:

- Has a valid DNS name that is resolvable.
- Does not have Chef server or client software installed on it.

About this task

In this task, you create an installation configuration of your solution by dragging objects onto a canvas, adding components to the objects, and defining properties for the components. The configuration defines the following solution nodes:

- Rapid Analytics node.
- Business Intelligence node.
- A data node.

Procedure

1. In a web browser, go to the Solution Installer.
2. Read the license agreement and click **Accept**.
3. Click **New Configuration** and then click **OK**.
4. Define the Business Intelligence node:
 - a. Under **Mandatory Content List**, drag a **Node** object onto the canvas.
 - b. Under **Property Editor**, enter values for the properties of the object. Set the **Name** field to indicate that this object is the Business Intelligence node.
 - c. Drag **Cognos BI Content** onto the Business Intelligence node object. Do not change the values under **Property Editor**.
 - d. Drag the **Cognos BI Report Images** object onto **Cognos BI Content** object. Do not change the values under **Property Editor**.
5. Define the Rapid Analytics node:
 - a. Under **Mandatory Content List**, drag a **Node** object onto the canvas.
 - b. Under **Property Editor**, enter values for the properties of the object. Set the **Name** field to indicate that this object is the Rapid Analytics node.
 - c. Drag the **Rapid Analytics Framework** object onto the Rapid Analytics node object. Do not change the values under **Property Editor**.
6. Define the data node:

- a. Under **Mandatory Content List**, drag a **Node** object onto the canvas.
 - b. Under **Property Editor**, enter values for the properties of the object. Set the **Name** field to indicate that this object is the data node.
 - c. Drag the **Database Content** object onto the data node object. Do not change the values under **Property Editor**.
7. If you want to save the configuration to use or refer to in the future, click **Save Configuration**.
 8. Click **Validate**. The progress of the validation is displayed. If the validation fails, look at the log file to troubleshoot the problem.
 9. Click **Run**. The scripts and files that are used to install and configure the solution are copied to the target nodes.

What to do next

Complete the installation of the solution on each of the nodes.

Installing the TELCO database

To install the TELCO database on the data node, you run a script that creates the TELCO database, tables, constraints, and indexes.

Procedure

1. Log on to the data node with a user ID that has access to the IBM DB2® database. For example, **db2inst1**.
2. Enter the following command to run the script:
db2 -tvf TELCO_Db2.sql
3. Verify that the database tables exist. Type the following commands:
 - a. db2 connect to db TELCO db2 list tables for schema NPS

Information similar to the following message is displayed:

FACTOR_IMPORTANCE	NPS	T	2015-06-07-22.46.41.072519
FACT_SCORE	NPS	T	2015-06-07-22.46.41.815746
MASTER_CUSTOMER	NPS	T	2015-06-07-22.46.41.833610
MASTER_DATA_ACTIVITY_LOCATION	NPS	T	2015-06-07-22.46.41.848970
MASTER_DEVICE	NPS	T	2015-06-07-22.46.41.864207

- b. db2 list tables for schema BBCI

Information similar to the following message is displayed:

CGR_DEVICE	BBCI	T	2015-09-09-02.12.13.907406
CHURN_DATA	BBCI	T	2015-08-20-15.56.39.964202
CHURN_MANAGEMENT	BBCI	T	2015-08-19-01.33.27.083252
CHURN_MANAGEMENT_HISTORY	BBCI	T	2015-09-11-16.47.45.466452
SUBSCRIBER_BILLING	BBCI	T	2015-09-11-16.03.21.410002
SUBSCRIBER_CARE	BBCI	T	2015-09-11-16.04.12.393064
SUBSCRIBER_CRM	BBCI	T	2015-09-11-16.05.00.176739
SUBSCRIBER_LEVEL_CX_SCORE_WEEKLY	BBCI	T	2015-08-20-16.02.13.77829

Deploying content on the Business Intelligence node

You must deploy the content files (report images and Framework Manager model) for the Behavior Based Customer Insight for Communication Service Providers solution.

Procedure

1. Log on to the BI node as the **root** user or as a user with **sudo** permissions.
2. Decompress the following files:

- CSP_CustomerInsight_FMProject.zip
 - CSP_CustomerInsight_Images.zip
3. Copy the CSP_CustomerInsight_BIContent.zip folder to the *Cognos_Install_location/Deployment* folder on the BI node computer.
 4. Copy the CSP_CustomerInsight_Images folder to the *Cognos_Install_location/webcontent/samples* folder on the BI node computer.
 5. Catalog the TELCO database on the computer where you installed Framework Manager. The following steps are optional but are a best practice:
 - a. Open a command prompt by selecting **Windows Start > Run**.
 - b. Type the following command:
db2cmd
 - c. Enter the following commands to catalog the database node command line editor:
catalog tcpip node *node_name* remote *dbserver_ip* server *port_number*
where *node_name* is the name of the computer where you installed Framework Manager and *dbserver_ip* is the IP address of the database server. The default *port_number* is 50000.
catalog database *database_name* at node *node_name* authentication server

For example,

```
catalog tcpip node TELCO remote 9.124.113.116 server 50000
catalog db TELCO at node TELCO authentication server
```

Creating a data source connection to the TELCO database

You must create a data source connection from the Business Intelligence node to the IBM DB2 database that is on the data node.

Procedure

1. Open a web browser on the BI node.
2. Go to the IBM Cognos BI portal URL. For example, go to `http:<hostname>/ibmcognos/`.
3. On the **Welcome** page, click **Administer IBM Cognos Content**.
4. Click the **Configuration** tab, and then click **Data Source Connections**.
5. Click **New Data Source**.
6. In the **Name** field, type TELCO, and then click **Next**.
7. From the **Type** list, select IBM DB2. Ensure that **Configure JDBC connection** is selected and click **Next**.
8. Based on the connection information that is provided by your database administrator, enter the appropriate information. If applicable, select the authentication mode to be used. Configure the signon credentials if applicable, and click **Next**.
9. On the **Specify the DB2 (JDBC) connection string** page, enter the server name, port number, and database name.
10. Click **Test the connection**, then click **Test**. On the **View the results** page, the status of the connection tests for the dynamic query mode is **Succeeded**.
11. Click **Close**, click **Close** again, and then click **Finish**.

Setting up the Cognos Framework Manager model

Open the project file in IBM Cognos Framework Manager and save it.

Before you begin

The user ID that you log in with when you open Framework Manager must belong to the **db2admins** group.

Procedure

1. On the Business Intelligence node, open Cognos Framework Manager.
2. Click **Open a project**.
3. Select the *File_saved_location/CSP_CustomerInsight_FMModel/CSP_CustomerInsight_FMModel.cpf* file.
4. Click **OK**. The model opens and you can view the data model, sources, and package.
5. Publish the package.

Deploying the IBM Cognos BI reports

You use IBM Cognos Administration to deploy the reports.

Before you begin

Ensure that *CSP_CustomerInsight_BIContent.zip* is in the *Cognos_install_location/Deployment* folder on the Business Intelligence node.

Procedure

1. Open a web browser on the Business Intelligence node.
2. Go to the IBM Cognos BI portal URL. For example, go to <http://cognosnode/ibmcognos/>.
3. On the **Welcome** page, click **Administer IBM Cognos Content**.
4. Click the **Configuration** tab, and then click **Content Administration**.
5. On the toolbar, click **New Import**.
6. In the **Deployment Archive** pane, select deployment archive (*CSP_CustomerInsight_BIContent.zip*), and click **Next**.
7. In the **Specify a name and description** pane, accept the default or enter a new name, and click **Next**.
8. In the **Select the public folders and directory content** pane, select all of the packages in the table, leave the **Options** as default, ensure that the **Disable after import** option is unselected, and click **Next**.
9. In the **Specify the general options** pane, accept the defaults, and click **Next**.
10. On the **Review the summary** page, click **Next**.
11. On the **Select an action** page, select **Save and run once**, and click **Finish**.
12. On the **Run with options** page, accept the defaults, and click **Run**, and then click **OK**.
13. Select **View the details of this import after closing this window** and click **OK**.
14. In IBM Cognos Administration, click **Home**. The content is available in **Public Folders > CSP_CustomerInsight_BIContent_Model > CSP_CustomerInsight > CSP_CustomerInsight_Analysis_Report**.

Deploying IBM SPSS content

You must create a data source connection, and then import the content files into IBM SPSS Deployment Manager for Collaboration and Deployment Services.

Before you begin

Ensure that the following things are in place:

- IBM SPSS Collaboration and Deployment Services Deployment Manager is installed on the client computer.
- The `CSP_CustomerInsight_CDS.pes` file is on the local client computer from which the deployment is performed. The default installation folder for this file is `/opt/IBM/IS_CustomerInsight_1.0/Analytics/`.
- You know the admin credentials for the Predictive Analytics node, IBM DB2, and IBM SPSS Collaboration and Deployment Services.

Procedure

1. Start **Deployment Manager**.
2. Click **File > New > Content Server Connection**.
3. In the **Connection Name** field, enter a name that identifies the Predictive Analytics node.
4. In the **Server URL** field, enter `http://analytics_node_IP:9080` and click **Finish**.
5. Right-click **Content Repository** and click **Import**.
6. Browse to and select the `CSP_CustomerInsight_CDS.pes` file.
7. Click **Open**.
8. Accept the default options in the **Import** window, and click **OK**. The **CSP_CustomerInsights_NPS** folder, a job in the **Jobs** folder, and streams in the **Modeler Streams** folder are created.
9. Update the credential information for the **Admin**, **Db2inst1**, and **root** users so that the job and streams run successfully.
 - a. In the **Content Explorer** tab, under **Resource Definitions**, open the **Credentials** folder.
 - b. Update the **Admin** user with the credentials for the SPSS Collaboration and Deployment Services user that has access to the content repository and runs the job.
 - c. Update the **Db2inst1** user with the credentials for the user that has access to the IBM DB2 database.
 - d. Update the **root** user with the credentials for the user that has access to the Modeler Server.
10. Open the job in the **Jobs** folder and ensure that the user credentials match the credentials of the IBM DB2 instance on the **ODBC** tab.
11. Create an ODBC connection with the name **IBMPCI** that points to the **TELCO** database on the data node.

For information about how to create an ODBC connection, see *Configuring ODBC for IBM SPSS Modeler Server on Linux operating systems*. In the `odbc.ini` file, specify the driver path as `/opt/IBM/SDAP71/lib/XEdb227.so`.

SPSS job sequencing and triggering

Specific processing jobs define execution sequences and triggering events for individual files stored in the IBM SPSS Collaboration and Deployment Services Repository.

For churn prediction, the jobs which perform the training of the model and the scoring of the model have been separated. The data for training needs to be prepared and loaded in the Behavior Based Customer Insight for Communication Service Providers churn tables. Once this is performed, the job `\CSP_CustomerInsight_NPS\Jobs\Training Churn Prediction Model` present in IBM SPSS Collaboration and Deployment Services needs to be run from the SPSS CADS Deployment Manager. This will create the predictive model nuggets which will be used subsequently in the job used to score the active customers. Failure to run this job before running the `\CSP_CustomerInsight_NPS\Jobs\Scoring Churn Prediction Model` will result in error.

The job `\CSP_CustomerInsight_NPS\Jobs\Scoring Churn Prediction Model` will be triggered each time the sqoop job for transferring the churn dataset from the rapid analytics node to the DB2 data node is complete.

The job `\CSP_CustomerInsight_NPS\Jobs\Load Report Tables` will be triggered each time the sqoop job for transferring the NPS dataset from the rapid analytics node to the DB2 data node is complete.

Changing churn factor labels

You can change the churn factor labels to suit your business needs.

Each time the SPSS training job is run, it creates a decision list model. The decision list model nugget is then exported to a .csv file so that it can be provided as an additional input apart from the churn propensity and displayed in the Churn Management Report. The idea is to provide actionable insights for the churn along with providing the churn propensity and the expected churn date.

The decision list usually has a list of a few variables expressed as an if/else condition. A business analyst or user can interpret these segments and provide labels which can be used in the churn management report. This will help the users of the churn management report understand the reason for the churn. Providing of labels is optional. In the event that the label is not provided, the base ruleset that appeared in the decision list will appear in the churn factors column.

To provide the **Churn_Factors** with easy to understand labels, the following steps need to be performed.

1. Copy the file `Decision List Formatted.csv` located in the folder `/usr/IBM/SPSS/ModelerServer/16.0/data` in the SPSS/Analytics node.
2. Open the comma separated file in a spreadsheet. You will see these columns: **ID**, **NBCORRECT**, **RECORD_COUNT SCORE**, **CHURN_RULES**, and **CHURN_FACTORS**.
3. Input values in the **Churn_Factors** column can be changed to something more meaningful. For instance, a churn rule from the decision list which states **OVERALL_EXP_SCORE_5 greaterThan 0.500000** and **OVERALL_EXP_SCORE_5 lessOrEqual 3** can be labeled as **Low Overall Experience Score** and entered in the **Churn_Factors** column.

4. In the example in step 3, the scoring stream will identify **Low Overall Experience Score** as the churn factor and display the same based on the Rule Id.

Improving the SPSS model performance

You can fine-tune the SPSS model performance.

The Model Training stream Churn Prediction Model Training.str is used to train the two predictive models that are used for scoring. The first predictive model is a decision list, which provides the rules to identify segments that have a higher propensity to churn, along with the propensity to churn. The second predictive model is a COX regression model, which uses the attributes of the customer to identify when they are likely to churn.

In the decision list, changing the following options can alter the performance characteristics, such as time needed to build the model and overall accuracy of the mode:

- Decreasing the minimum segment size will create smaller segments. When decreasing the minimum segment size, the maximum number of segments need to be increased correspondingly, so that you can adequately cover the population. For instance, with a minimum segment size of 0.5% and a maximum number of groups of 100, we are likely to cover upwards of 50% of the total population based on their propensity to churn. The rest of the population can be considered as likely to be active. However, increasing the maximum number of segments is likely to increase the model development time proportionately.
- Increasing the confidence interval for new conditions will decrease the number of segments formed. However, they will be more reliable segments.

The analysis weight is currently set as 3 for churned customers and 1 for active customers. This is set in the supernode **Data Cleansing** in the node **Analysis Weight**. This results in oversampling of churned customers. This will help provide importance to identifying churned customers. If the number of active customers is very high and there is a need to decrease model development time, the analysis weight can be changed to 1 for churned customers and 0.34 for active customers. This will result in undersampling of active customers and provide improvements in the model development time.

The **Analysis Weight** proportion should reflect the business priorities of how important it is to identify customers likely to churn over the ones that remain active. In an environment where there is a customer churn of 25% a year, an analysis weight of 3:1 might be good to make the importance equal in proportion. In an environment where customer churn is already 50%, an analysis weight proportion of 1:1 might be sufficient.

Balancing can be used effectively to reduce the number of records instead of analysis weight, which could reduce the time needed to train for both the Decision List & Cox Regression Model Development.

The time required for **Model Scoring** Churn Prediction Model Scoring.str can be improved by having intermediate tables from which the data can be loaded. Subsequent sections of the stream can read the data for scoring. This can be used to have SQL pushback features increase the speed of the stream overall by distributing a significant portion of the workload to the database.

Splitting the data for scoring purposes into multiple divisions and running the streams in parallel can help improve the performance.

The Rapid Analytics platform

The Rapid Analytics platform consists of the framework layer and the mediation layer.

In the Behavior Based Customer Insight for Communication Service Providers solution, the Rapid Analytics node contains the framework layer of the Rapid Analytics platform. The Mediation node contains the mediation layer of the Rapid Analytics platform.

The mediation layer uses a component called `vedia` to collect, process, and aggregate data. The data is stored in IBM InfoSphere BigInsights. The framework layer contains IBM InfoSphere BigInsights.

The Rapid Analytics platform components are in the following folders:

- `rapidanalytics/build/ra-platform`
- `rapidanalytics/build/vedia_analytics`
- `rapidanalytics/build/vedia_gn`
- `rapidanalytics/build/vedia_lte`
- `rapidanalytics/build/vedia_voice`

In the context of the Behavior Based Customer Insight for Communication Service Providers solution, the BigInsights master node is the Rapid Analytics node.

Installing the framework layer

You install the framework layer on the Rapid Analytics node.

About this task

In the context of the Behavior Based Customer Insight for Communication Service Providers solution, the BigInsights master node is the Rapid Analytics node.

Procedure

1. Log in to the Rapid Analytics node as the **root** or as a user with **sudo** permissions.
2. Copy the Rapid Analytics platform tarball to the `/opt/tnf/install` folder.
3. Decompress the tarball.
4. Change to the directory containing the platform rpm:

```
cd /opt/tnf/install/rapidanalytics/build/ra-platform
```
5. If a previous installation of Rapid Analytics platform exists, remove it. Type the following command:

```
rpm -e bis-main
```
6. Install the RPM file. Type the following command:

```
rpm -i bis-main-2*.rpm
```
7. Switch to a BigInsights administrator user. For example, type the following command:

```
su - biadmin
```
8. Go to main platform installation directory. Type the following command:

```
cd /opt/tnf/apps/bis-main
```

9. Verify that the hadoop, catalog, and bigsql components are running by typing the following commands:

```
status.sh hadoop
status.sh catalog
status.sh bigsql
```

10. If the components are not running, type the following command to start all of them:

```
start.sh
```

To start individual components, you can use the `start.sh` command with the components to start.

Configuring the Rapid Analytics node

You must configure the properties in the Rapid Analytics platform and the files in the platform.

About this task

There are three configuration files in the `apps/bis-main-var/bis-demon/cfg` folder on the Rapid Analytics node that you must review and if necessary, modify for your environment. Perform this procedure for each of the following files:

- `cfg.properties`
- `bigsql_connection.properties`
- `db2_connection.properties`

Refer to the following table for the properties to modify in the `cfg.properties` file:

Table 1. Properties in the `cfg.properties` file

Property	Description	Default value	Required?
<code>db.schema.name</code>	Hive schema name	TNF	Yes
<code>use_bigsql_for_aggregations</code>	Uses Big SQL for aggregation queries, otherwise Hive is used.	false	No
<code>segments_by_imsi</code>	Number of segments by IMSI for TDRs and customer level aggregation tables. This must be same value that is in all vedia configurations.	10	Yes
<code>min_5_aggregation.keep.days</code>	Number of days the system keeps data of 5 minute aggregation tables.	8	Yes
<code>h_1_aggregation.keep.days</code>	Number of days the system keeps data of hourly aggregation tables.	40	Yes
<code>d_1_aggregation.keep.days</code>	Number of days the system keeps data of daily aggregation tables.	90	Yes
<code>tdr.tmp_min_10.keep.days</code>	Number of days the system keeps data of 10 minute temporary TDR tables.	1	Yes

Table 1. Properties in the `cfg.properties` file (continued)

Property	Description	Default value	Required?
<code>tdr.keep.day</code>	Number of days the system keeps data of TDR tables.	40	Yes

Refer to the following table for the properties to modify in the `bigsql_connection.properties` file:

Table 2. Properties in the `bigsql_connection.properties` file

Property	Description
<code>db.url</code>	JDBC URL connection string to the IBM InfoSphere BigInsights database. For example, <code>jdbc:db2://localhost:51000/bigsql</code> . Update the port number if your database uses a port that is different from the default.
<code>db.user</code>	Database user ID
<code>db.password</code>	Database user password

Refer to the following table for the properties to modify in the `db2_connection.properties` file:

Table 3. Properties in the `db2_connections.properties` file

Property	Description
<code>db.url</code>	JDBC URL connection string to the IBM DB2 database. For example, <code>jdbc:db2://localhost:50000/tnfbis</code> . Update the port number if your database uses a port that is different from the default.
<code>db.user</code>	Database user ID
<code>db.password</code>	Database user password

Procedure

1. Log on to the Rapid Analytics node as the **biadmin** user.
2. Open a configuration file, review the properties, and modify as required. Refer to the tables for information on the properties to update.

Configuring table groups

You can configure table groups before or after you start the Rapid Analytics platform server.

About this task

The tables are configured after you start the Rapid Analytics platform server so you may not see them immediately.

The tables in each table group are listed in the following table:

Table 4. Tables in each table group

Table group file name	Tables in the group
gn-subscribers-aggregation.tablegroup	gncontrolp_3_h_1 gncontrolp_3_min_5 gncontrolpst_3_h_1 gncontrolpst_3_min_5 gnuserp_st_2_h_1 gnuserp_st_2_min_5 gnuserplane_3_h_1 gnuserplane_3_min_5
gn-tdrs.tablegroup	tdr6013 tmp6013 tdr6021 tmp6021 tdr6022 tmp6022
lte-subscribers-aggregation.tablegroup	ltes11_3_h_1 ltes11_3_min_5 ltes11_st_3_h_1 ltes11_st_3_min_5 ltes1u_3_h_1 ltes1u_3_min_5 ltes1u_st_2_h_1 ltes1u_st_2_min_5
lte-tdrs.tablegroup	tdr7013 tmp7013 tdr7021 tmp7021 tdr7022 tmp7022
voice_sms-subscribers-aggregation.tablegroup	voice_sms_3_h_1 voice_sms_3_min_5 voice_sms_sc_3_h_1 voice_sms_sc_3_min_5

Table 4. Tables in each table group (continued)

Table group file name	Tables in the group
voice_sms-tdrs	tdr_voicesms
	tmp_voicesms

Procedure

1. Log in to the Rapid Analytics node and type the following command:
`cd /opt/tnf/apps/bis-main/bis-dbconfig`
2. Configure the table groups that are required for your environment. For example, to configure the **gn-subscribers-aggregation** table group, type the following command:
`./install-db.sh gn-subscribers-aggregation`
To get a complete list of available and configured table groups, type the following command: `./install-db.sh`

What to do next

Start the Rapid Analytics platform server.

Starting the Rapid Analytics platform server

Start the daemon server in the framework layer.

Procedure

1. Log in to the Rapid Analytics node as the **biadmin** user.
2. Type the following commands:
`cd /opt/tnf/apps/bis-main`
`./bis-main.sh start`
3. Confirm that the server started successfully by looking at the following log file:
`/opt/tnf/apps/bis-main/bis-demon/log/demon.log`
4. Deploying datasets. Once the Rapid Analytics Server has been started, follow the instructions in the following documents found in the install directory `/opt/IBM/IS_CSP_CustomerAnalytics_1.1/rapidanalytics/build` or `/opt/IBM/IS_CSP_CustomerInsights_1.1/rapidanalytics/build`:
 - *Net-Promoter-Score Readme.pdf*
 - *Customer-Behaviors-Readme.pdf*
 - *Customer-Experience-Readme.pdf*
 - *Churn-Readme.pdf*
5. Follow the instructions in the document *CSP CustomerInsight DBLoader Setup document.pdf*.

Installing the SLM tagging functionality

You can deploy the slm tagging functionality on the rapid analytics node.

About this task

A RPM Package Manager (RPM) is contained in the `rapidanalytics/build/slm/ci-slm-deployment-1.0.0.rpm` folder. To install this RPM, follow these steps:

Procedure

1. Install the SLM via the RPM installer by typing the following command:

```
rpm -i bis-main-2*.rpm
```

The files will be copied to `/opt/tnf/apps/slm`.

2. Running SLM tagging:

- A Cron job needs to be created that runs the SLM tagging command daily, just after midnight:

```
java -cp /opt/tnf/apps/slm/lib/log4j-1.2.16.jar:/opt/tnf/apps/slm/lib/* com.tnf.slm.SlmMain
```

- The SLM tagging process appends to a log file in the directory, with the extension `.slmtag`: `/opt/tnf/apps/slm/log/<log file>.slmtag`

This log file is used by IBM to ensure that the product is being used according to its license.

Installation and configuration of the Mediation node

Vedia is the component that processes the transaction data records (TDRs) on the Mediation node.

You install two versions of the vedia component on the Mediation node. Vedia analytics collects transaction data record (TDR) data, batches it, and loads it into flat files in IBM InfoSphere BigInsights. Vedia rapid calculates counters and metrics, groups the TDRs by key field values, and loads the aggregated data into IBM InfoSphere BigInsights.

Installing the mediation layer

Data collection is done by two versions of vedia mediation on the Mediation node.

Before you begin

The computer that is the Mediation node must:

- Have a `biadmin` UNIX account created on it.
- Have an `/opt/tnf` directory and the `biadmin` account must have write access to it.
- Have an `/var/opt/tnf` directory and the `biadmin` account must have write access to it.

About this task

In this task you install the vedia mediation, the Data Loader tool, and the IBM InfoSphere BigInsights client on the Mediation node. Perform this procedure on every instance of the Mediation node.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the `/etc/security/limits.conf` file and add the following lines or modify existing settings to these values:

```
biadmin hard nofile 655365
biadmin soft nofile 65536
biadmin hard nproc 65536
biadmin soft nproc 65536
```

- Open a web browser and open the BigInsights web console by typing the following address:

`http://bi-master-hostname:8080`

where *bi-master-hostname* is the IP address or hostname of the Rapid Analytics node.

Tip: If you see the message “Unable to connect” in your web browser, verify that the BigInsights web console is running on the Rapid Analytics node by using the `status.sh` console command. If the BigInsights web console is not running, start it with the `start.sh` console command.

- Log in as **biadmin**.
- Under **Quick Links**, click **Download client library and development software**.
- Click **Download** beside **InfoSphere BigInsights client libraries**.

Tip: If you do not see **InfoSphere BigInsights client libraries**, verify that the `rpm-build` package is installed on the Rapid Analytics node by typing the following command: `sudo yum install rpm-build`.

The `bi-client.rpm` file is downloaded to your computer.

- Copy the `bi-client.rpm` file from your computer to the `/opt/tnf/install` directory on the Mediation node.
- Type the following command:


```
rpm -i bis-client.rpm
```
- Decompress the `CSP_CustomerInsight_RA.zip` file into `/opt/tnf/install`.
- Type the following commands:


```
cd /opt/tnf/install/rapidanalytics/build/ra-platform
rpm -i bis-main-*.rpm
```

Installing the mediation components

You must install the vedia mediation components. There is one component for each network type.

About this task

The following table shows the distribution packages for each network type:

Table 5. Distribution packages for each network type

Vedia type	File name	Description
Rapid Gn	<code>install_vedia_rapid_1.0.0_gn.sh</code>	Install script
	<code>vedia_rapid_1.0.0_gn_config.tar.gz</code>	Configuration package
	<code>vedia_rapid_1.0.0_gn_bin.tar.gz</code>	Binaries package
Rapid Lte	<code>install_vedia_rapid_1.0.0_lte.sh</code>	Install script
	<code>vedia_rapid_1.0.0_lte_config.tar.gz</code>	Configuration package
	<code>vedia_rapid_1.0.0_lte_bin.tar.gz</code>	Binaries package
Rapid Voice and SMS	<code>install_vedia_rapid_1.0.0_voice.sh</code>	Install script
	<code>vedia_rapid_1.0.0_voice_config.tar.gz</code>	Configuration package
	<code>vedia_rapid_1.0.0_voice_bin.tar.gz</code>	Binaries package
Vedia Analytics (raw TDRs)	<code>install_vedia_rapid_8.2.1.sh</code>	Install script

Table 5. Distribution packages for each network type (continued)

Vedia type	File name	Description
	vedia_rapid_8.2.1_config.tar.gz	Configuration package
	vedia_rapid_8.2.1_bin.tar.gz	Binaries package

The following table shows the component and installation directories:

Table 6. Component and installation directories for each network type

Project	Component directory	Default installation directory (\$INSTALL_DIR)
Rapid Gn	rapidanalytics/build/vedia_gn	/opt/tnf/apps/rapid_gn
Rapid Lte	rapidanalytics/build/vedia_lte	/opt/tnf/apps/rapid_lte
Rapid Voice & SMS	rapidanalytics/build/vedia_voice	/opt/tnf/apps/rapid_voice
Analytics	rapidanalytics/build/vedia_analytics	/opt/tnf/analytics/vedia

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Decompress the distribution packages and installation script in the `/opt/tnf/install` directory.
3. Change to the relevant install directory for the media component that you are installing and run the corresponding shell script: For example, to install the vedia mediation component for the *Gn* network type, type the following commands:

```
cd /rapidanalytics/build/vedia_gn
sh install_vedia_rapid_1.0.0_gn.sh
```

4. Install all four vedia mediation components.

Configure the vedia mediation components

Vedia is the component that processes and aggregates the transaction data records (TDRs) on the Mediation node. You must configure the probe collectors, probe access, probe directories, and other settings in the configuration files in the `INSTALL_DIR/cfg` directory.

Configuring IMSI segments on the Mediation node

Configure the number of IMSI segments in the `const_db.xml` file to match the number in the `cfg.properties` file on the Rapid Analytics node.

About this task

For information about configuring the `cfg.properties` file, see “Configuring the Rapid Analytics node” on page 15.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the `consts-db.xml` file.

- Set the **value** parameter in the following line to the number of IMSI segments that are in your environment:

```
<const name="hadoop-segments" value="1" description="Hadoop segments"/>
```

Set this value in the consts-db.xml file on all instances of the Mediation node.

Configuring the Gn and LTE probe collectors

About this task

Use the following information to configure the attributes in the VediaMain.xml file:

- Specify all collectors for the same probe within one application attribute.
- Specify unique values for each occurrence of the probe-no attribute in the file.
- The value of the probe-name attribute is used in the **Collector** dimension in the database tables.

Procedure

- Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
- Open the VediaMain.xml file and if necessary, modify it for your environment.

In the following example, one probe **P1** has collector **ip1** and another probe **P2** has collectors **ip2** and **ip3**. **ip1**, **ip2**, and **ip3** are the IP addresses of the probes:

```
<vedia name="Main config">
  <hosts description="Main host and Collectors hosts">
    <host name="Collector" address="localhost" main="1"/>
  </hosts>

  <!-- One collector with name P1, define by probe-name="P1", it can have some probes with
        diff probe-no=".." -->
  <applications name="Collector1" host="Collector" probe-name="P1">
    <application name="Collection_1" config="Collection" probe-ip="ip1" probe-no="1"/>
  </applications>

  <!-- Next collector define with name P2 (probe-name="P2"),
        all probe-no=".." and name=".." should be different in this config -->
  <applications name="Collector2" host="Collector" probe-name="P2">
    <application name="Collection_2" config="Collection" probe-ip="ip2" probe-no="2"/>
    <application name="Collection_3" config="Collection" probe-ip="ip3" probe-no="3"/>
  </applications>
</vedia>
```

- Configure the probe collection directories, user name, and password in /cfg/consts.xml.

The following example shows the configuration of the probe collection directories, user name, and password:

```
<const name="probe-user" value="tnf" description="Ftp user"/>
  <const name="probe-password" value="tnf" description="Ftp password"/>
  <const name="probe-dir" value="/apps/midas/bin/MANAGE/process/dacas/netrofeed/"
description="Ftp dir"/>
not expected to be changed <const name="probe-files_mask"
value="netro_{[0-9]{14}}\.log\.gz" description="Ftp files mask"/>
not expected to be changed <const name="probe-file-date-format"
value="yyyyMMddHHmmss" description="Ftp files mask"/>

  <const name="probe-pb-on" value="no"
description="Enable/Disable page based records (yes/no)"/>
  <const name="probe-pb_dir" value="/apps/midas/bin/MANAGE/process/dacas/pagebasedqoe/"
description="Ftp dir"/>
not expected to be changed <const name="probe-pb_files_mask"
value="pbqoe_{[0-9]{14}}\.log\.gz" description="Ftp files mask"/>
not expected to be changed <const name="probe-pb_file-date-format"
```

```

value="yyyyMMdHHmmss" description="Ftp files mask"/>

<const name="vedia-number" value="01"
description="Unique number for each instance of Vedia"/>
<!-- Probe constants -->
<!-- Roaming Type calculation
0 = "home" - to be used for Gn and LTE probes
1 = "roamer" - to be used for Gp probes
2 = "roam_home" - ideally should not happen -
will be used in case if records from different interfaces (Gn and Gp for instance)
are populated in a single feed.
-->
<const name="probe-traffic-type" value="0"
description="Probe traffic type (0 = home/1 = roamer/2 = roam_home)"/>
<const name="probe-mcc" value="272" description="Probe MCC"/>
<const name="probe-mnc" value="07" description="Probe MNC"/>

```

Configuring the rapid voice and SMS probe

If you installed the rapid voice and SMS project type, then you must configure the VediaMain.xml and consts.xml files.

About this task

Use the following information to configure the attributes in the VediaMain.xml file:

- Specify a user ID and password for the user that accesses the probe ftp or sftp server.
- Specify the directory that contains the probe data.

Use the following information to configure the attributes in the consts.xml file:

- Specify all collectors for the same probe within one application attribute.
- Specify unique values for each occurrence of the probe-no attribute in the file.
- The value of the probe-name attribute is used in the **Collector** dimension in the database tables.
- The probe-type attribute defines the type of interface on the probe. The valid values for it are **a** for an A interface and **iucs** for an IuCS interface.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the VediaMain.xml file and if necessary, modify it for your environment.

In the following example there is one probe **P1** that has an A interface and another probe **P2** that has an IuCS interface:

```

<vedia name="Main config">
  <hosts>
    <host name="Collector" address="localhost" main="1"/>
  </hosts>

  <applications name="Collector1" host="Collector" probe-name="P1" probe-type="a">
    <application name="Collection_1" config="Collection" probe-ip="ip1" probe-no="1"/>
  </applications>

  <applications name="Collector2" host="Collector" probe-name="P2" probe-type="iucs">
    <application name="Collection_2" config="Collection" probe-ip="ip2" probe-no="2"/>
    <application name="Collection_3" config="Collection" probe-ip="ip3" probe-no="3"/>
  </applications>
</vedia>

```

3. Set the parameters for the two voice interface feeds in consts.xml. The following example shows the parameters:

```

<!-- Ftp loader for voice-sms A interface -->
  <const name="probe-a-user" value="cinnara" description="Ftp user"/>
  <const name="probe-a-password" value="cinnara" description="Ftp password"/>
  <const name="probe-a-dir" value="/var/tmp/data/bin/a" description="Ftp dir"/>
  <const name="probe-a-files_mask" value="Aif_ASN1_BER_[0-9]{1,3}\.[0-9]{1,3}\.
[0-9]{1,3}\.
[0-9]{1,3}_([0-9]{8}\.[0-9]{4,6})" description="Ftp files mask"/>
  <const name="probe-a-file-date-format" value="yyyyMMdd.HHmss"
description="Ftp files mask"/>
  <const name="probe-a-csv-dir" value="/var/tmp/data/csv/a"
description="Dir with decoded csv files"/>
  <const name="probe-a-csv-files_mask" value="Aif_([0-9]{8}\.[0-9]{6})_[0-9]{5}
.csv.gz"
description="Csv file mask"/>

  <!-- Ftp loader for voice-sms IuCS interface -->
  <const name="probe-iucs-user" value="cinnara" description="Ftp user"/>
  <const name="probe-iucs-password" value="cinnara" description="Ftp password"/>
  <const name="probe-iucs-dir" value="/var/tmp/data/bin/iucs" description="Ftp dir"/>
  <const name="probe-iucs-files_mask" value="IuCS_ASN1_BER_[0-9]{1,3}\.[0-9]{1,3}
\.[0-9]{1,3}\.
[0-9]{1,3}_([0-9]{8}\.[0-9]{4,6})" description="Ftp files mask"/>
  <const name="probe-iucs-file-date-format" value="yyyyMMdd.HHmss"
description="Ftp files mask"/>
  <const name="probe-iucs-csv-dir" value="/var/tmp/data/csv/iucs"
description="Dir with decoded csv files"/>
  <const name="probe-iucs-csv-files_mask" value="IuCS_([0-9]{8}\.[0-9]{6})_[0-9]{5}
.csv.gz"
description="Csv file mask"/>

```

Configuring the probe traffic type

The probe traffic type applies to only the Gn and LTE standards. The traffic type values are used to calculate roaming dimensions such as Home MCC and MNC, Visited MCC and MNC, and roaming type.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the `consts.xml` file.
3. Update the values as shown in the following example:

```

<const name="probe-traffic-type" value="0" description="Probe traffic type
(0 = home/1 = roamer/2 = roam_home)"/>

<const name="probe-mcc" value=".." description="Probe MCC"/>
<const name="probe-mnc" value=".." description="Probe MNC"/>

```

where `probe-traffic-type=0` – "home" - use for Gn and LTE probes, `probe-traffic-type=1` – "roamer" - use for Gp probes, `probe-traffic-type=2` – "roam_home" - should not happen. Use if records from different interfaces (Gn and Gp for instance) are populated in a single feed, `probe-mcc=` – Default probe MCC, `probe-mnc=` – Default probe MNC.

4. Open the `mcc-mnc.csv` file and modify it for your environment. For example:

```

#MCC|MNC|OPERATOR TYPE: 0 = international, 1 = home,
2 = probe|Country|Country Code|Network
289|88|0|Abkhazia|7|A-Mobile
289|68|0|Abkhazia|7|A-Mobile

```

Configuring a roamers filter

You configure a roamers filter in the `consts.xml` file.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the `consts.xml` file.
3. Modify the values for the roamers as shown in the following example:

```
<!-- Roamers filter
subscriber - Roamers IN subscriber level - IMSI is used as aggregation key, MSISDN
and IMEI fields are populated
operator - Roamers IN operator level - MCC/MNC extracted from IMSI is used as an
aggregation key instead of IMSI. The rest of IMSI characters are replaced by zeros.
MSISDN and IMEI fields are not populated.
filter - Roamers IN are filtered out. TDRs generated by roamers in are filtered out and
not used for aggregations
-->
<const name="roamers-filter" value="subscriber" description="Roamers filter
(subscriber/operator/filter)"/>
```

where `roamers-filter=subscriber` is the default value. Roamers IN subscriber level - IMSI is used as aggregation key, MSISDN and IMEI fields are populated
`roamers-filter=operator` – Roamers IN operator level - MCC/MNC extracted from IMSI is used as an aggregation key instead of IMSI. The rest of IMSI characters are replaced by zeros. MSISDN and IMEI fields are not populated
`roamers-filter=filter` – TDRs where IMSI's MCCMNC is not provisioned as local are filtered out and not used for aggregations.

Configuring a cells filter

You configure a cells filter in the `consts.xml` file.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open the `consts.xml` file.
3. Modify the values for the roamers as shown in the following example:

```
<!-- Check probe cells in cells.csv.
Only cell populated in the provisioning file should be taken into account and CI/LAC
fields be used for aggregations.
If cell is not provisioned in the file, then CI/LAC value should be replaced by zero
-->
<const name="check-probe-cells" value="no" description="Check probe cells in cells.csv
file (yes/no)"/>
```

where `check-probe-cells=no` means disable cells filter by default,
`check-probe-cells=yes` means enable cells filter. The cells specified in the `cfg/cells.csv` file are accounted for and CI/LAC fields are used for aggregations. Cells that re not provisioned are aggregated as **Uncategorized**.

Here is an example of the `cfg/cells.csv` file:

```
#CellLac|CellID
12295|10891
12295|10985
12295|14098
12295|14181
```

Configuring the Rapid Analytics platform for TDR processing

If you want the Behavior Based Customer Insight for Communication Service Providers solution to process raw transaction data records (TDRs) you must update some configuration files.

About this task

The configuration files for TDR processing are stored in `$INSTALL_DIR/cfg/`.

The following table list the parameters in the consts.xml file that you must change or verify:

Parameter	Description
hadoop-segments	Number of IBM InfoSphere BigInsights segments. This value must match the number of segments configured in IBM InfoSphere BigInsights.
wait-min	Wait time before sending an hourly .csv file to IBM InfoSphere BigInsights.
vedia_col.host	Host name of the collector
vedia-number	Number that is appended to the file names uploaded to IBM InfoSphere BigInsights. This value must be different in each vedia that is processing TDRs for the same interface.

The following table list the probe properties that you must change or verify:

Parameter	Description
probe-user	User name of the probe FTP server.
probe-password	Password of the probe FTP server.
probe-file-date-format	Ftp files mask.
probe_type="NT"	Gn probe.
probe_type="LTE"	S1U-S11 probe.
probe_type="PB"	Gn page based (6013).
probe_type="LTE_PB"	LTE page based (7013).

The VediaMain.xml contains collection information for each probe. The following table shows the parameters for a probe:

Parameter	Description
name	Name of the probe.
config	Collection.
probe-type	NT for GN probe LTE for S1U-S11 probe PB for GN page-based probe LTE_PB for LTE page-based probe
probe-no	Probe number. This value must be different for each collector in the same blade.
description	Description of the probe.

Procedure

1. Log on to Mediation node as the **root** user or as a user with **sudo** permissions.
2. Open and verify or edit each file.

Loading data and starting the predictive analytics job

After the Rapid Analytics platform is installed, you must configure and run the data load process and start the predictive analytics job.

Before you begin

Ensure that the following conditions are met:

- The Rapid Analytics platform is installed.
- You have a user ID that has access to the IBM InfoSphere BigInsights database that is on the Rapid Analytics node.
- IBM InfoSphere BigInsights is running.
- The IBM DB2 database exists on the data node.

About this task

The `runDBExport.sh` script extracts data from the IBM InfoSphere BigInsights database on the Rapid Analytics node and loads it into the IBM DB2 database on the data node. The script takes a parameter that runs the following variations of the script:

runDBExport.sh 1

Extracts and loads the entire database. Run this script to do the initial extract and load of the data.

runDBExport.sh 2

Extracts and loads the data that has changed (delta) between the last time the extract and load process ran and now. Run this script for subsequent extract and load processing.

The following table lists the parameters in the `config.properties` file that you must configure:

Table 7. Parameters in the config.properties file

Parameter	Description	Default value
HiveHost	Change the hive host and port number	localhost:10000/default
HiveUser	Hive user name	hive
HivePasswd	Hive user password	No default value.
DB2Host	DB2 Host and DB2 port number	localhost:50000
DB2User	DB2 user name	db2inst1
DB2Passwd	DB2 user password	No default value.
SqoopHost	Change the Sqoop host address and sqoop port number	localhost:12000/sqoop
HiveDataDirectory	Location of the Hive warehouse directory on the Rapid Analytics node	/biginsights/hive/warehouse/ tnf.db
SPSSEndPoint	URL of the IBM Predictive Customer Intelligence job. Update the host address.	http://localhost:9080/ process/services/ ProcessManagement
SPSSID	Job ID of SPSS. The OBJECT URI of the job to be started. You can find this value by right-clicking the job in Deployment Manager and selecting properties.	No default value.
SPSSUser	SPSS CADS User ID. Used to log in to the Deployment Manager.	admin
SPSSPasswd	SPSS CADS Password.	No default value.

Procedure

1. Log on to the Rapid Analytics node.
2. Change directory to the location of the CSP_CustomerInsight_RA.zip solution package.
3. Decompress the IS_CSP_CustomerInsight_1.0/RapidAnalytics/net-promoter-score/CSP_CustomerInsight_DB.zip file.
4. Copy the following files to the *SqoopHomeDir/lib* and *HiveHomeDir/lib* folders:
 - db2jcc.jar
 - db2jcc_licence_cu.jar
5. Open the config.properties file and change the parameters to refer to IBM InfoSphere BigInsights and the IBM DB2 database. Refer to the table in *About this task* for information about the parameters to change.
6. Enter the following commands:

```
chmod 755 runDBExport.sh
chmod 755 script/sqoop.sh
```
7. Go to the IS_CSP_CustomerInsight_1.0/RapidAnalytics/net-promoter-score/Start CSP_CustomerInsight directory and type the following command:

```
./runDBExport.sh 1
```
8. Run CSP_CustomerInsight for every new load of the InfoSphere BigInsights database by using the following command:

```
./runDBExport.sh 2
```

Chapter 3. Using Behavior Based Customer Insight for Communication Service Providers

This section describes a typical user case that shows how to use the Behavior Based Customer Insight for Communication Service Providers solution to understand customer experience trends.

IBM Behavior Based Customer Insights for Communication Service Providers is composed of three separate components:

- a net promotor score (NPS) correlation application
- a churn management application
- a customer behavior dataset

The NPS correlation and churn management applications include an IBM Cognos user interface, which are described in the workflow procedures.

The customer behavior dataset provides the following:

- a set of tables customers can access
- an accompanying data dictionary

The churn management application provides an analysis of subscriber data suggesting how likely and when a customer is likely to churn.

The dataset contains subscribers' usage data, including their interests and preferences. The application analyzes the amount of time subscribers spend browsing, searching, reading, and streaming information on their devices. Each period of activity is categorized, where possible, into one of hundreds of interest categories, enabling communication service providers to categorize their subscribers into hundreds of different micro-segments based upon these interests.

Workflow and user roles

This workflow shows how the NPS correlation application provides a dashboard and key metrics to a communication service provider. For this example, a customer experience team member investigates the NPS score for mobile data. A typical process workflow is described.

The workflow also provides a set of key metrics demonstrating the churn management application. For this example, the likelihood of a subscriber's churn propensity is associated with their churn propensity score. This functionality gives communication service providers (CSPs) the ability to identify customers who are likely to churn and also those of greatest value to the CSP. A typical process workflow is described.

Using the NPS correlation application

The Behavior Based Customer Insight for Communication Service Providers solution provides in-depth analytics to help communication service providers evaluate key subscriber metrics.

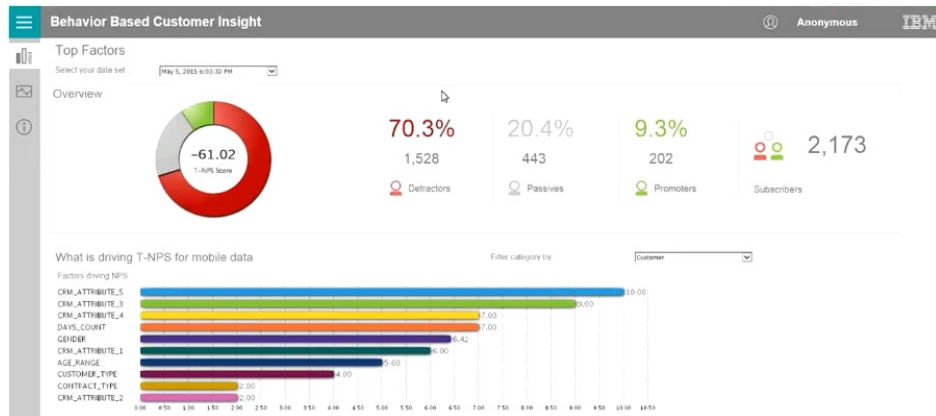
About this task

This task describes how a typical customer experience team member would use the reports to investigate net promoter score (NPS) and transactional-NPS. The data is linked to the actual activity of subscribers. Subscriber data such as services, devices, or location are linked to their NPS, providing an enhanced view of the network and services.

The screen shots shown are for example purposes and will vary depending on your particular implementation.

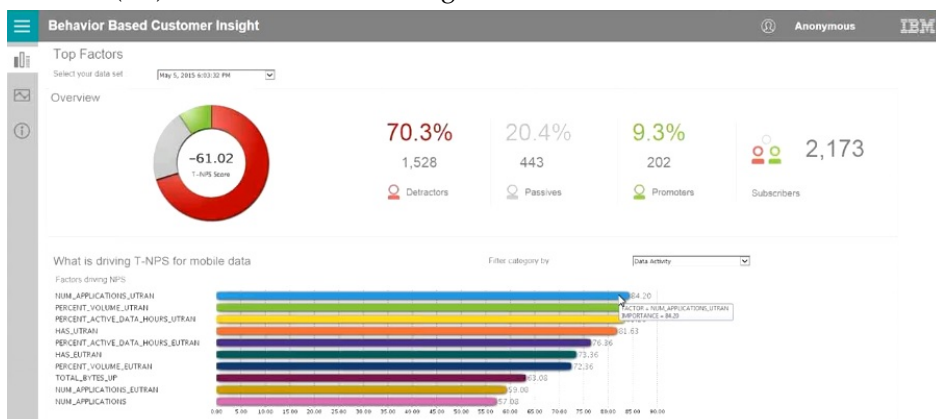
Procedure


1. Open the **Behavior Based Customer Insight Top Factors** dashboard.

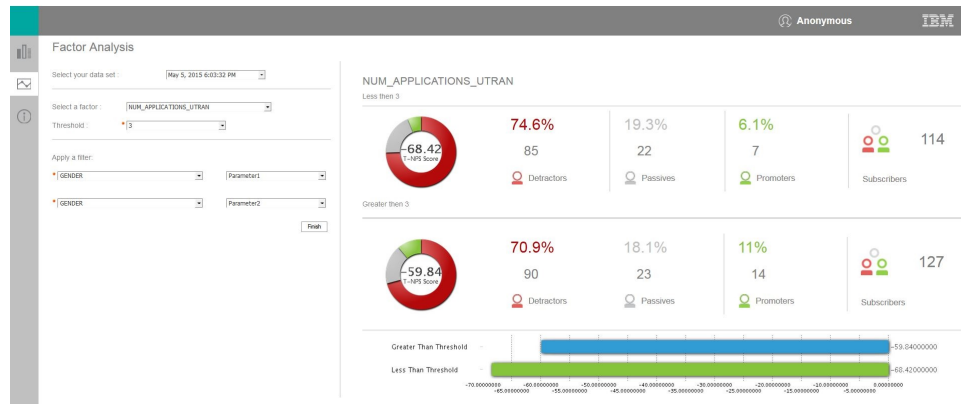


The T-NPS score of -61.02 indicates a high number of unhappy subscribers. This number reflects promoters (approximately 9%) minus detractors (approximately 70%). Factors driving the NPS data are summarized in the bar chart and legend.

2. Filter the data category based on **Data Activity**. The top 10 factors change to reflect the **Data Activity** category. The number of applications used on the UTRAN (3G) network is the most significant.

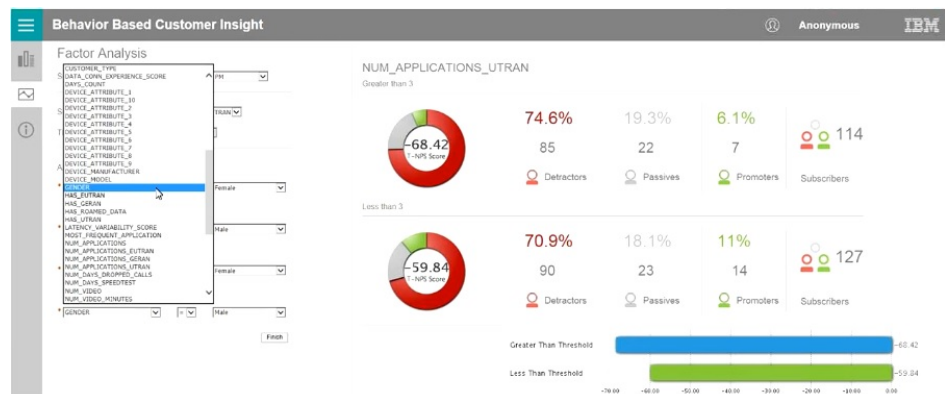


- Click the Factor Analysis icon  to launch the **Factor Analysis** page for NUM_APPLICATIONS_UTRAN.



With the Factor Analysis threshold set at 3, subscribers who are running more than three applications on UTRAN (a 3G network) have much lower NPS scores than those who are using three or less. With a similar number of subscribers, this analysis is statistically significant. The customer experience team can then use this as a means of further investigation.

- Apply further filters to the analysis, setting the data to show only female customers.



With this filter applied, the analysis shows an even greater difference between females who use more than three applications on 3G versus females who use three applications or less on 3G. A customer experience team member might then use this information for further investigation. For example, what type of applications were involved in creating this apparent negative experience for the subscribers in question?

Using the churn management application

The churn management application utilizes standard churn management inputs, as well as voice, SMS, and mobile data records, to analyze trends in subscriber voice, SMS, and mobile data and to provide an interpretation of the churn model.

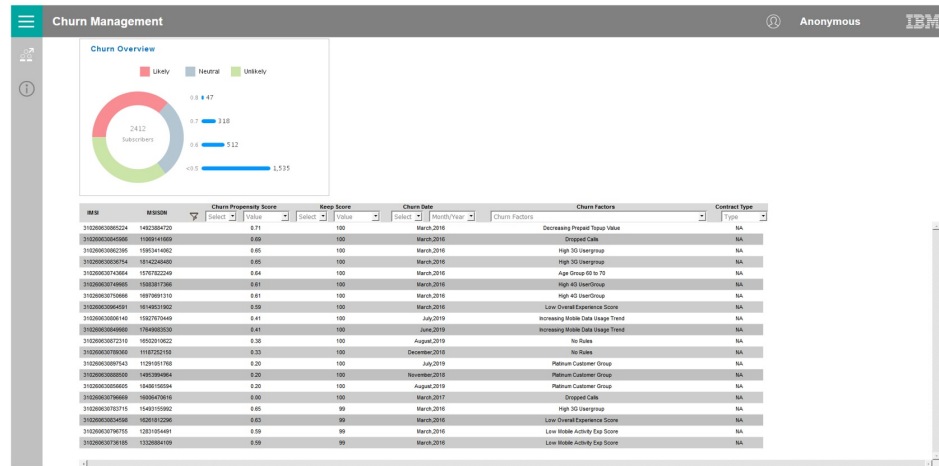
About this task

This task describes how a typical customer experience team member would use the reports to investigate a set of subscribers and evaluate their likelihood to churn based on their churn propensity and other factors.

The screen shots shown are for example purposes and will vary depending on your particular implementation.

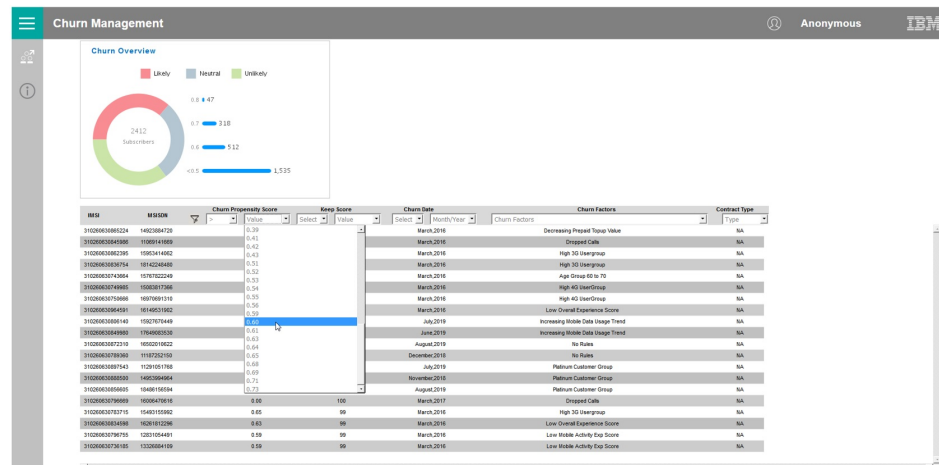
Procedure

1. Open the **Churn Management** report.

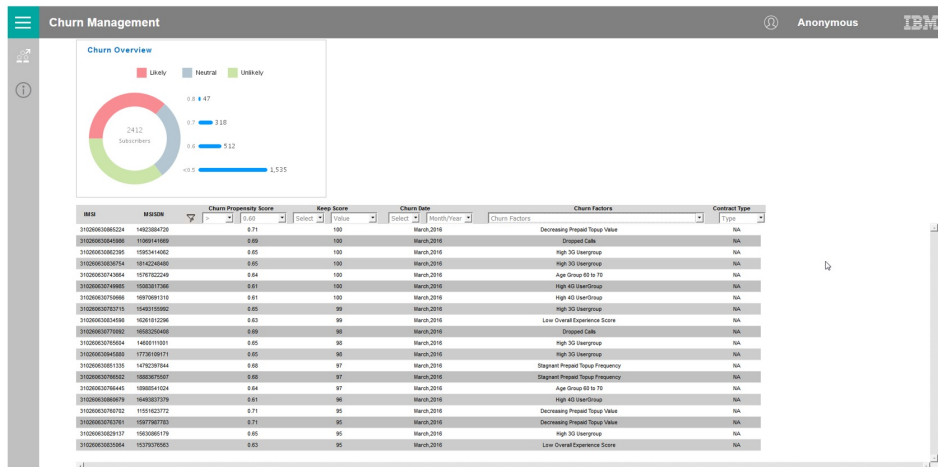


The main report page shows an overview of the entire subscriber base in the donut chart. Of the 2412 subscribers, 877 are likely to churn based on their churn propensity of ≥ 0.6 . The default report is sorted first by **Keep Score** and then **Churn Propensity Score** (both descending). This highlights to communication service providers (CSPs) those subscribers who are the greatest value, as well as those who are most likely to churn.

2. Using the pull-down menus on the report, select the greater than sign ($>$) and then select a value of **0.60**.

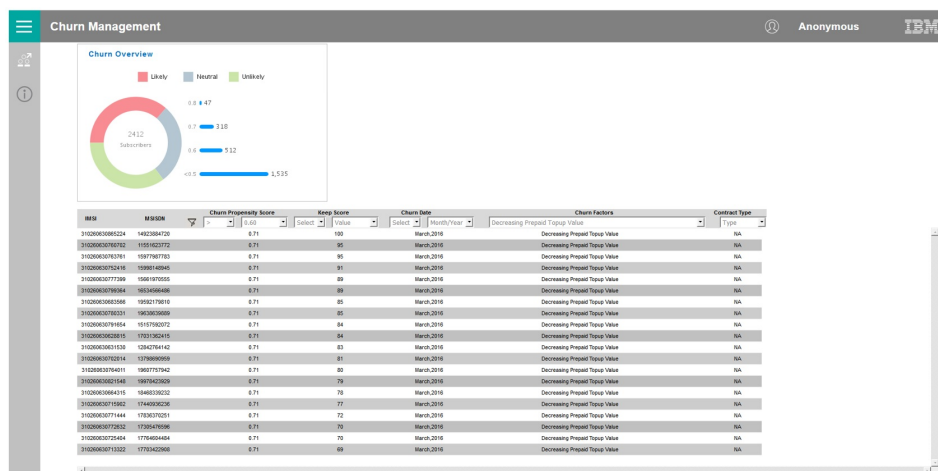


This filters the subscribers in the **Churn Propensity Score** column to include only those with a score greater than 0.60.



Four other factors are available for filtering:

- **Keep Score:** Filters by relative subscriber importance. For example, this factor would allow a customer experience team member to concentrate on those subscribers with a score greater than 50.
 - **Churn Date:** Provides a projected churn date. For example, this filter would allow a customer experience team member to concentrate on those likely to churn sooner rather than later
 - **Churn Factor:** Provides a suggested churn date. This filter would allow a customer experience team member to segment based on specific factors impacting subscriber satisfaction, including offers which might be available to likely churners (such as **Decreasing Prepaid Top Value**).
 - **Contract Type:** Filters by the type of subscribers. For example, this filter would allow a customer experience team member to segment based on **Prepay** or **Postpay** subscribers, which may be assigned to different departments
3. Click the **Churn Factors** pull-down menu, and select **Decreasing Prepaid Top Value**.



With this filter applied, subscribers are listed whose predominant reason for churn is the decreasing value of their monthly top-up, ranked in descending order by **Keep Score** and **Churn Propensity Score**. In this example, there are

several very valuable subscribers who are likely to churn by March 2016. These are the subscribers who the CSP would most likely address when attempting to stop subscribers from churning.

Appendix A. Accessibility features

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products.

For information about the commitment that IBM has to accessibility, see the IBM Accessibility Center (www.ibm.com/able).

IBM Cognos HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.

Report output

In IBM Cognos Administration, you can enable system-wide settings to create accessible report output. For more information, see the *IBM Cognos Business Intelligence Administration and Security Guide*. In IBM Cognos Report Studio, you can enable settings to create accessible output for individual reports. For more information, see the *IBM Cognos Report Studio User Guide*. You can access the previously mentioned documents at IBM Knowledge Center (<http://www.ibm.com/support/knowledgecenter>).

Appendix B. The data provision tool

The data provision tool loads, deletes, and exports data from the provision tables. The provision tables contain customer relationship management data.

The data provision tool is installed in the `/opt/tnf/apps/bis-main-var/bis-provisioning-tool` folder and consists of the following parts:

load.sh

Loads and updates provision tables.

load-all.sh

Checks all the files in the `.csv` folder and loads them at the same time.

delete.sh

Reads records from input files and deletes records from IBM InfoSphere BigInsights by using key values.

export.sh

Exports records from IBM InfoSphere BigInsights to text files.

Usage

You must specify key values for all operations. The key field names are obtained by the provision tools from the table metadata. The primary key must be defined in table DDL. In the following examples the primary key field is **tac**:

```
create table device (  
tac varchar(32) primary key, model varchar(64),  
vendor varchar (64)  
);
```

The follow example contains two primary keys:

```
create table device (  
tac varchar(32),  
model varchar(64), vendor varchar (64), primary key (tac, model)  
);
```

To see the parameters for `load.sh` type, the script name with no parameters in a command-line window:

```
Usage: load.sh [options] Options:  
* --file, -f  
CSV file with data * --table, -t  
DB Table name
```

--file, -f

A list of input files that are separated by the pipe (`|`) character.

--table, -t

The table name. The table name can include a dot separated schema. For example, `load.sh -f device.csv -t device`.

`delete.sh` and `export.sh` take the same parameters.

The data provision tool reads table metadata and converts the input text file into an Apache Parquet or plain text formatted file and populates specified table with the data. It updates existing records and inserts new records.

Creating dictionary tables

Run the `/opt/tnf/apps/bis-main/bis-dbconfig/install-db.sh` script with table group name parameters. The script `/opt/tnf/apps/bis-main/bis-main.sh` creates a dictionary table in the database when it starts.

You cannot modify the data in IBM InfoSphere BigInsights directly. The data provision tool updates the database tables and exports them to IBM InfoSphere BigInsights. During the export stage, the data provision tool must know the storage format. Apache Parquet and plain text formats are supported. Update the storage format in the `/opt/tnf/apps/bis-main/bis-tools/bisprovisioning-tool/cfg/hive-storage.properties` file.

Input files

The data input format is a pipe (`|`) separated utf-8 encoded plain text file with headers. The header can start with a number sign (`#`). Header values must match the dictionary table case insensitive field names. For example,

```
TAC|MODEL|VENDOR
12345678|Zamyang|Ciaobao
```

Configuration files

The data provision tool IBM DB2 database configuration location is `/opt/tnf/apps/bis-main-var/bis-demon/cfg/db2_connection.properties`. For example,

```
db.url=jdbc:db2://localhost:50000/tnfbis
db.user=catalog
db.password=db2
```

where `db.url` is the JDBC URL string of the database, `db.user` is the database user ID, and `db.password` is the database password.

Appendix C. Troubleshooting

This section contains troubleshooting information for the Behavior Based Customer Insight for Communication Service Providers installation.

Troubleshooting a problem

Troubleshooting is a systematic approach to solving a problem. The goal of troubleshooting is to determine why something does not work as expected and how to resolve the problem.

Review the following table to help you or customer support resolve a problem.

Table 8. Troubleshooting actions and descriptions

Actions	Description
A product fix might be available to resolve your problem.	Apply all known fix packs, or service levels, or program temporary fixes (PTF).
Look up error messages by selecting the product from the IBM Support Portal, and then typing the error message code into the Search support box (http://www.ibm.com/support/entry/portal/).	Error messages give important information to help you identify the component that is causing the problem.
Reproduce the problem to ensure that it is not just a simple error.	If samples are available with the product, you might try to reproduce the problem by using the sample data.
Ensure that the installation successfully finished.	The installation location must contain the appropriate file structure and the file permissions. For example, if the product requires write access to log files, ensure that the directory has the correct permission.
Review all relevant documentation, including release notes, technotes, and proven practices documentation.	Search the IBM Knowledge Center to determine whether your problem is known, has a workaround, or if it is already resolved and documented.
Review recent changes in your computing environment.	Sometimes installing new software might cause compatibility issues.

If the items in the table did not guide you to a resolution, you might need to collect diagnostic data. This data is necessary for an IBM technical-support representative to effectively troubleshoot and assist you in resolving the problem. You can also collect diagnostic data and analyze it yourself.

Troubleshooting resources

Troubleshooting resources are sources of information that can help you resolve a problem that you are having with an IBM product.

Support Portal

The IBM Support Portal is a unified, centralized view of all technical support tools and information for all IBM systems, software, and services.

The IBM Support Portal lets you access all the IBM support resources from one place. You can tailor the pages to focus on the information and resources that you need for problem prevention and faster problem resolution. Familiarize yourself with the IBM Support Portal by viewing the demo videos (https://www.ibm.com/blogs/SPNA/entry/the_ibm_support_portal_videos).

Find the content that you need by selecting your products from the IBM Support Portal (<http://www.ibm.com/support/entry/portal>).

Before contacting IBM Support, you will need to collect diagnostic data (system information, symptoms, log files, traces, and so on) that is required to resolve a problem. Gathering this information will help to familiarize you with the troubleshooting process and save you time.

Service request

Service requests are also known as Problem Management Reports (PMRs). Several methods exist to submit diagnostic information to IBM Software Technical Support.

To open a PMR or to exchange information with technical support, view the IBM Software Support Exchanging information with Technical Support page (<http://www.ibm.com/software/support/exchangeinfo.html>).

Fix Central

Fix Central provides fixes and updates for your system's software, hardware, and operating system.

Use the pull-down menu to navigate to your product fixes on Fix Central (<http://www.ibm.com/systems/support/fixes/en/fixcentral/help/getstarted.html>). You may also want to view Fix Central help.

IBM developerWorks

IBM developerWorks[®] provides verified technical information in specific technology environments.

As a troubleshooting resource, developerWorks provides easy access to the most popular practices, in addition to videos and other information: developerWorks (<http://www.ibm.com/developerworks>).

IBM Redbooks

IBM Redbooks[®] are developed and published by the IBM International Technical Support Organization, the ITSO.

IBM Redbooks (<http://www.redbooks.ibm.com>) provide in-depth guidance about such topics as installation and configuration and solution implementation.

Software support and RSS feeds

IBM Software Support RSS feeds are a quick, easy, and lightweight format for monitoring new content added to websites.

After you download an RSS reader or browser plug-in, you can subscribe to IBM product feeds at IBM Software Support RSS feeds (<https://www.ibm.com/software/support/rss>).

Log files

Log files can help you troubleshoot problems by recording the activities that take place when you work with a product.

Error messages

The first indication of a problem is often an error message. Error messages contain information that can be helpful in determining the cause of a problem.

Troubleshooting known data provision tool error messages

There are solutions for known data provision tool error messages that you may encounter in IBM Behavior Based Customer Insight for Communication Service Providers.

Table 9. Solutions to known data provision tool error messages in IBM Behavior Based Customer Insight for Communication Service Providers.

Error message	Action
<p>Permission denied.</p> <p>This error is a permission denied error with one of the following two messages:</p> <ol style="list-style-type: none"> 1. java.io.IOException. Cannot run program "runDBExport.sh": error=13, Permission denied. 2. java.io.IOException. Cannot run program "script/sqoop.sh": error=13, Permission denied. 	<p>Investigate the DBExport.log to search for the error.</p> <ol style="list-style-type: none"> 1. Check the permission of runDBExport.sh. Refer to the setup documentation for the appropriate steps. 2. Check the permission of sqoop.sh in the script folder in which the sqoop.sh resides. Refer to the setup documentation for the appropriate steps.
<p>Sqoop job failure</p> <ol style="list-style-type: none"> 1. IOException [main] LoadDB2 ERROR - java.io.IOException: Can't export data, please check task tracker logs. ERROR tool.ExportTool: Error during export: Export job failed! 2. IllegalArgumentException. [main] LoadDB2 ERROR - Caused by: java.lang.IllegalArgumentException: Timestamp format must be yyyy-mm-dd hh:mm:ss[.fffffffff] 	<ol style="list-style-type: none"> 1. Verify whether the hive data from the nps_score_table (if it's option 1) or the nps_score_table_latest (if it's option 2) contains the new line character (\n) at the EOF. Tip: The location of the hive data is specified by the HiveDataDirectory located in the config.properties file. 2. Verify whether the hive data from the nps_score_table (if it's option 1) or the nps_score_table_latest (if it's option 2) contains valid data in the field mentioned by the value of HiveDataDirectory located in the config.properties file.
<p>SPSS job error</p> <p>ERROR - Exception !!!</p> <p>java.io.IOException: Server returned HTTP response code: 500 for URL: http://<SPSSHost>:<port>/process/services/ProcessManagement</p>	<p>Verify whether SPSS is running and the correct host name and port number have been used.</p>

Troubleshooting the FTP collection of files in Rapid Analytics

There are solutions for known issues in the FTP collection of files in Rapid Analytics.

On Rapid Analytics mediation nodes, the function in Rapid Analytics that collects transaction data record (TDR) files for storage in Rapid Analytics is Vedia.

To verify that the collection of TDRs from the CSP probe or probes is working, note that each probe's collector will be defined in mediation. Use the secure shell (SSH) to connect to the correct collector. The allocation of collectors can be found in the file VediaMain.cfg at /opt/tnf/apps/vedia/cfg/VediaMain.cfg.

To check FileLoader logs, the format is logs/Collection_1/FileLoader092708.log.

When the application is running, messages would appear in the following format. For this example, Vedia is comparing TDRs in tmp against files on the probe:

```
[CFileListJob::GetNewFilesList] [TaskTest] Getting list of /var/opt/tnf/tdrs/ *.*
[CFileListJob::GetNewFilesList] [TaskTest] Files: 196, Match: 182, Old: 178
```

The files that are being collected correctly are shown in the following example:

```
[CFileDownloadJob::Process] Task - downloading midas/netrofeed/netro_20130918200012.log.gz
[CFileDownloadJob::Process] Task - download netro_20130918200012.log.gz (9600142)
[CProbe::Store] Save: /opt/tnf/apps/vedia/tmp/Collection_1_FileLoader.dat
[CFileListJob::Save] LastListTime: 20130918200012, Files: 664
[CFileListJob::GetNewFilesList] [Task] Getting list of midas/netrofeed/ *.*
[CFileListJob::GetNewFilesList] [Task] Files: 251, Match: 223, Old: 221
[CFileDownloadJob::Process] Task - downloading midas/netrofeed/netro_20130918200015.log.gz
[CFileDownloadJob::Process] Task - download netro_20130918200015.log.gz (9013063)
```

In contrast, the following is an example of a file that can no longer connect to a probe. Typically, if FTP is not working, the ports are not open, permissions are not set up correctly on the probe, the probe is down, or connectivity between the probe and mediation has dropped.

FTP not working

TO FIND example of ftp failing

Full queues represent full applications within the chain. The following is an example of a file that has full queues somewhere along the chain. If one application is full, then the preceding application cannot output to the queue, thus impacting the queues (creating a knock-on effect). The queues will need be examined for skipping.

```
1C8F - 19/09 11:00:55.200 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:00.242 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C91 - 19/09 11:01:00.461 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:05.402 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:10.431 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:15.482 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:20.642 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:25.838 - [CFileListJob::GetNewFilesList]
Skip getting of files list
```



```
1C91 - 19/09 11:01:30.525 - [CFileListJob::GetNewFilesList]
Skip getting of files list
1C8F - 19/09 11:01:31.063 - [CFileListJob::GetNewFilesList]
Skip getting of files list
```

If skipping is noticed in Vedia logs and does not stop, restart Vedia. Some skipping will occur on restart, but wait up to five minutes for it to clear. It typically takes one to five minutes for the skipping messages to clear from the logs.

Recommendations:

- When collecting files and **Match** is **0**, the regex in the FileLoader configuration is incorrect. Review and correct this in the configuration steps.
- If FileLoader understands the files from the network to be **Old**, contact the person responsible for the production of network files to correct the date and time of the files.
- Occasionally, FTP jobs will fail to collect a file, but FileLoader will retry. Investigate situations where there are many failures, causing files that fail to be collected. Check the configuration and verify that the mediation node can reach the host of the network files (probe). If there is a network problem, contact the person responsible for the network files.
- When the configuration is correct and network connectivity is confirmed, restart the Vedia processes and ensure there are no further reported exceptions.

Queues commonly will fill after a long outage period. Vedia will attempt to collect all files and process them immediately. The number of files is determined by a period of time and is set in `Collection.xml`. If the look back period is, for example, greater than 20 minutes and there are **Skip getting of files list** messages in the log, then reset the look back period. Carefully tune anything greater than 20. Do not simply increase to 300 minutes, as this can overload Vedia.

The following is configured in `Collection.xml`:

```
module name="FileLoader" id="FileLoader">
  <probe name="FileLoader" log-file="FileLoader" log--process="2">
    <files>
      <ftp id="Task" host="Data"
        url="ftp://{probe-ip}" port="21"
        user="{probe-user}" password="{probe-password}"
        remotedir="{probe-dir}" storagedir="tmp/data_{probe-no}"
        filename-regex="{probe-files_mask}"
        ftp-timeout="60" ftp-downloads="1" attempt-timeout="1"
        delete-processed="0" process-file-ignore-age="0"
        max-process-file-age="20" time-shift="-3600"
        compression="GNUZIP"
        start-format="{probe-file-date-format}"
        detection-interval="60">
      </ftp>
    </files>
    <output name="LoadedFiles" model="File" limit="50"/>
  </probe>
</module>
```

`process-file-ignore-age` directs Vedia to load all the old files. Always set this to **0** in a production environment.

`max-process-file-age` defines, in minutes, how old the files on the probe can be before they will be rejected as being too old.

`detection-interval` indicates how often to check for files, in seconds.

download-threads indicates how many threads to download with. Note that, if this is set cumulatively across all collectors to be a higher value than the probe can process, this may impact the probe and it may reject files.

max-try indicates how many times to attempt to download a file.

Troubleshooting SPSS

If your SPSS job fails, there are some troubleshooting steps you can follow to resolve the issue.

Procedure

1. Right-click on the **Load Report Tables** job in the **Jobs** folder of CSP_CustomerInsights_NPS.
2. Click **Show Job History**. This will display information about the job runs, their logs, and their status.
3. Select the job that has failed. The job will have three individual SPSS streams. Select the one which has failed.
4. Open the log associated with the failed stream.
5. Scroll to the end of the job to discover the reason the job failed.

Results

Common reasons for the job to fail include the unavailability of resources such as:

- Database
- Memory
- Any of the prerequisite services such as SPSS Modeler Server and CADS

After running the solution installer, Chef processes are still running after executing the cleanup.sh script

After running the cleanup.sh script on the solution installer node, all chef processes should be terminated. However, in some cases chef processes are still running.

If you try to run the solution installer setup.sh file again, you may receive a message indicating that required ports are in use, or creation of the chef user might fail.

The solution to this issue is to reboot the installation node after having run "cleanup.sh".

Note: This only applies to the solution installer node and not the client/target nodes.

Notices

This information was developed for products and services offered worldwide.

This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. This document may describe products, services, or features that are not included in the Program or license entitlement that you have purchased.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Software Group
Attention: Licensing
3755 Riverside Dr.
Ottawa, ON
K1V 1B7
Canada

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

Trademarks

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “ Copyright and trademark information ” at www.ibm.com/legal/copytrade.shtml.

The following terms are trademarks or registered trademarks of other companies:

- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- UNIX is a registered trademark of The Open Group in the United States and other countries.
- Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Index

A

accessibility 35

B

BigInsights
 web console 19

C

cells filter 25
Churn Management application 31
churn prediction 13
Collaboration and Deployment Services Deployment
 Manager 11, 12
configuration 5
configuration files
 cfg.properties 21
 Collection.xml 42
 const.xml 22, 23, 24, 25
 db2_connection.properties 37
 Vediamain.cfg 42
 VediaMain.xml 22, 23, 25
customer relationship data 37

D

data load 27
data provision tables
 loading 37
data provision tool 37
 troubleshooting 41
data source connection
 ODBC 11
 TELCO database 9
database
 DB2 8
 TELCO 8
Deployment Manager 11, 12

F

filter
 cells 25
 roamers 25
Framework Manager 10

I

IBM SPSS 11, 13, 27
 churn factors 12
 model performance 13
 sequencing 12
 triggering 12
IMSI 21
installation 5
 DB2 8
 deploying installation files 7

installation (*continued*)
 Mediation node 19, 20
 overview 6
 prerequisites 5
 TELCO database 8

J

job
 configuring 11
 starting 27

M

master node 5
mediation 21
 See vedia component
Mediation node 19
 installing 19, 20

N

nodes
 Business Intelligence node 8
 Mediation 19, 20
 requirements 7
NPS correlation application 30

P

package
 contents 5
probe
 Gn 22, 24
 LTE 22, 24
 properties 25
 rapid voice and SMS 23

R

Rapid Analytics 14, 18
 configuring 15
 framework 14
 node 15
 transaction data record (TDR) 42
 troubleshooting 42
reports 10
roamers filter 25

S

servers
 See nodes
solution
 overview 1
 topology 5
starting 18

T

table groups 16
TDR 25
transaction data record (TDR) 25
troubleshooting 39
 data provision tool 41

V

vedia component 14, 19, 21, 25