National Broadband Map Data Assessment

Data Comparison Methodology

Data Comparison Methodology

Contents

General Methodology

Methodology

The wireline SBI data is compared to four sources of broadband information: FCC Speed Test, FCC Form 477, Centris, and Ookla. The wireless SBI data is compared to FCC Speed Test, FCC Form 477, Mosaik Solutions, and Ookla. Each SBI record that intersects a given census block is checked against each record in the comparison datasets that intersect that same census block. When the provider name matches, the technology of transmission code is checked. When the technology of transmission code also matches, the speed tier attributes are checked for matches.

The SBI broadband records include wireline and wireless providers and technologies. The wireline records include DSL, cable modem, optical fiber, electric power line, and other copper wireline technologies. The wireless broadband records include satellite, fixed wireless, and mobile wireless technologies. Each record is a unique provider name and technology combination and can represent an address point, road segment, census block, or wireless service area polygon.

Each of these records has six attributes that are compared to the FCC and commercial broadband datasets: provider name, technology of transmission, maximum advertised upstream speed, maximum advertised downstream speed, typical upstream speed, and typical downstream speed. When the SBI broadband data are compared to the FCC and commercial broadband datasets, the number of matches for each of the six attributes is recorded at the original record level (census block, road segment, address point, and wireless service area polygon).

The SBI data cannot be compared to the comparison datasets in native form. There is no common attribute to link the records. To perform the comparisons, all broadband datasets are associated with a common spatial feature. The census block was chosen to be the common spatial feature. When a data source does not have census block IDs, a census block ID is added to each record. For the SBI data, all address point records are assigned the ID number of the census block they intersect, and each road segment, and wireless service area are split into new features by each census block they intersect. Once all of the comparisons are performed, the matches for any road segment or wireless service area that has been split into multiple parts are added together and averaged, then assigned to the original single record.

At the census block level, a unique record is defined by the combination of provider name, technology code, and census block ID. This eliminates duplicate records but still allows for a

provider who provides access to multiple broadband technologies within the same census block. For each unique record, the match scores for each of the six broadband record descriptors are added. These descriptors are: provider name, technology of transmission, maximum advertised downstream speed, maximum advertised upstream speed, typical downstream speed, and typical upstream speed. The maximum match score equals the number of comparison data sources used in the comparison process. For all broadband records, the maximum match score for provider name and technology of transmission is 4. The maximum match score for the maximum advertised speed attributes is 3 and the maximum match score for the typical speed attributes is 2. Three of the comparison sources provide maximum speeds, and two provide typical speeds.

Comparison Results

The results of the comparisons are written to a results table, with "matches" recorded as 1, "no match" recorded as 0, and "unable to perform comparison" recorded as 9. An "unable to perform comparison" value is used under the following circumstances:

1. A 9 is populated in any comparison field (V_*) of the RLT table except V_PROVIDER_NAME, if one of the corresponding fields (either SBI data or comparison data) has a value of null, and a valid comparison cannot be made.
2. A 9 is populated in the four speed tier comparison fields (V_*) of the RLT table to show that a comparison was stopped because the provider name + census block ID for that record match but the technology values do not. When the technology values do not match, the speed tiers are not compared, to avoid invalid comparisons.
3. A 9 is populated in any of the four speed comparison fields (V_*) if a comparison source does not have information for that speed field.
4. A 9 is populated in V_PROVIDER_NAME if a part of the record being compared falls outside of the state boundary and does not intersect with a census block for the appropriate state.

Final Match Scores

Final results are appended onto the original SBI data (census blocks, road segments, address points, and wireless service areas) at the record level.

Description

The first step in creating the final match scores is the creation of a score accumulation table. A matched score count and a possible total score count are calculated for each record per comparison field. This presents a visual representation of the total number of matches for a record, for all comparison sources combined, compared to the total number of possible matches.

The road segment and wireless service area records in the accumulation table are still broken into parts of the original record by the census blocks they intersect. To assign the final match score to the original SBI record the match scores for the parts must be summed and then divided by the number of parts that make up the original record. The result is the final match score for each original record. The accumulation table records for SBI address point and census block records are simply summed for each unique SBI record ID to develop the final score for those records.

During the creation of the final match scores, nine's (9) are considered to be zero's (0). The final match score value domain is:

- 0 indicates no match

- > 0 and < or = 4 indicates the number of matches

- 7 indicates that the provider name for that SBI record was not found in any of the broadband comparison sources, so could not be compared.

Final matching scores are then attached to the original SBI record data to be used in the National Broadband Map. There are six match scores for each record:

1. Provider Name;

2. Technology of Transmission;

3. Maximum Advertised Download Speed;

4. Maximum Advertised Upload Speed;

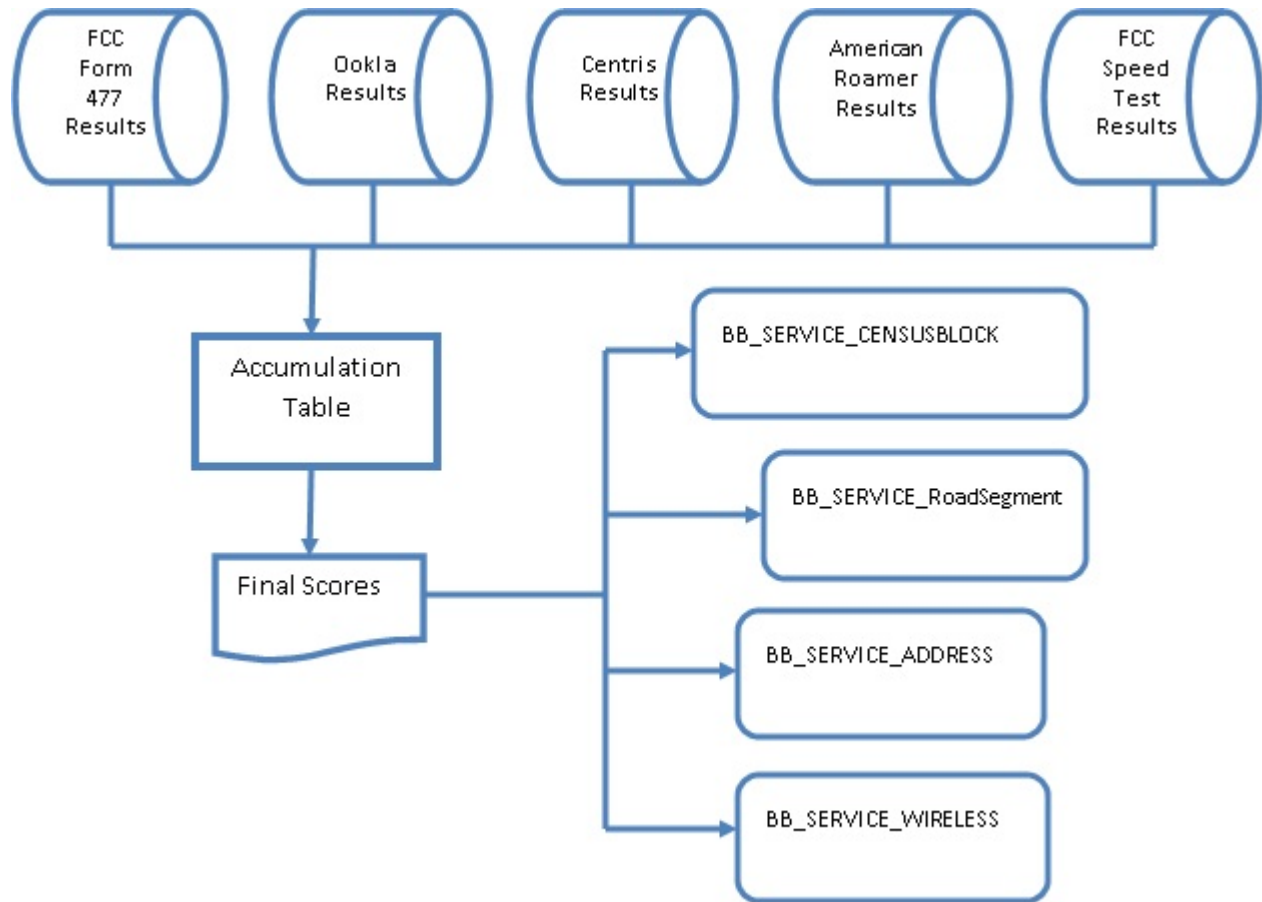5. Typical Download Speed;

6. Typical Upload Speed.

Process

Figure 1.   Final Scoring Process

Comparison Source Details

This section provides additional detail concerning each of the comparison data sources used to develop the match scores for each SBI update.

The wireline SBI data is compared to four sources of broadband information: FCC Speed Test, FCC Form 477, Centris, and Ookla. The wireless SBI data is compared to FCC Speed Test, FCC Form 477, American Roamer, and Ookla.

FCC Form 477

The FCC Form 477 data is information collected from service providers about broadband internet access connections to households and businesses. This is done in a data collection called Local Telephone Competition and Broadband Reporting (FCC Form 477). Information has been collected each year, as of June 30 and December 31, starting with December 31, 1999. This dataset is updated when a newly published version of the data is made available.

FCC Form 477 data is compared to the SBI Awardee-provided broadband service census blocks, road segments, address points, and wireless service areas. This comparison source contains wireline and wireless data and is used to compare provider names, technology of transmission, and maximum advertised speeds at the census tract level. Form 477 provider names, speed tiers, and technology of transmission codes differ from those in the SBI data, so lookup tables were created for matching purposes.

Comparison Source Preparation Description

FCC Form 477 data is already in a format similar to that of the SBI broadband data. The highest maximum advertised download and upload speed tiers are used for each combination of parent company name, technology, and census tract. Therefore, if any slower speed information exists in the Form 477 data for that combination of parent company name, technology, and census tract, it is not used for the data comparisons. The 477 records are then loaded into the state comparison database.

Comparison Fields

Table 1. FCC Form 477 Comparison Fields

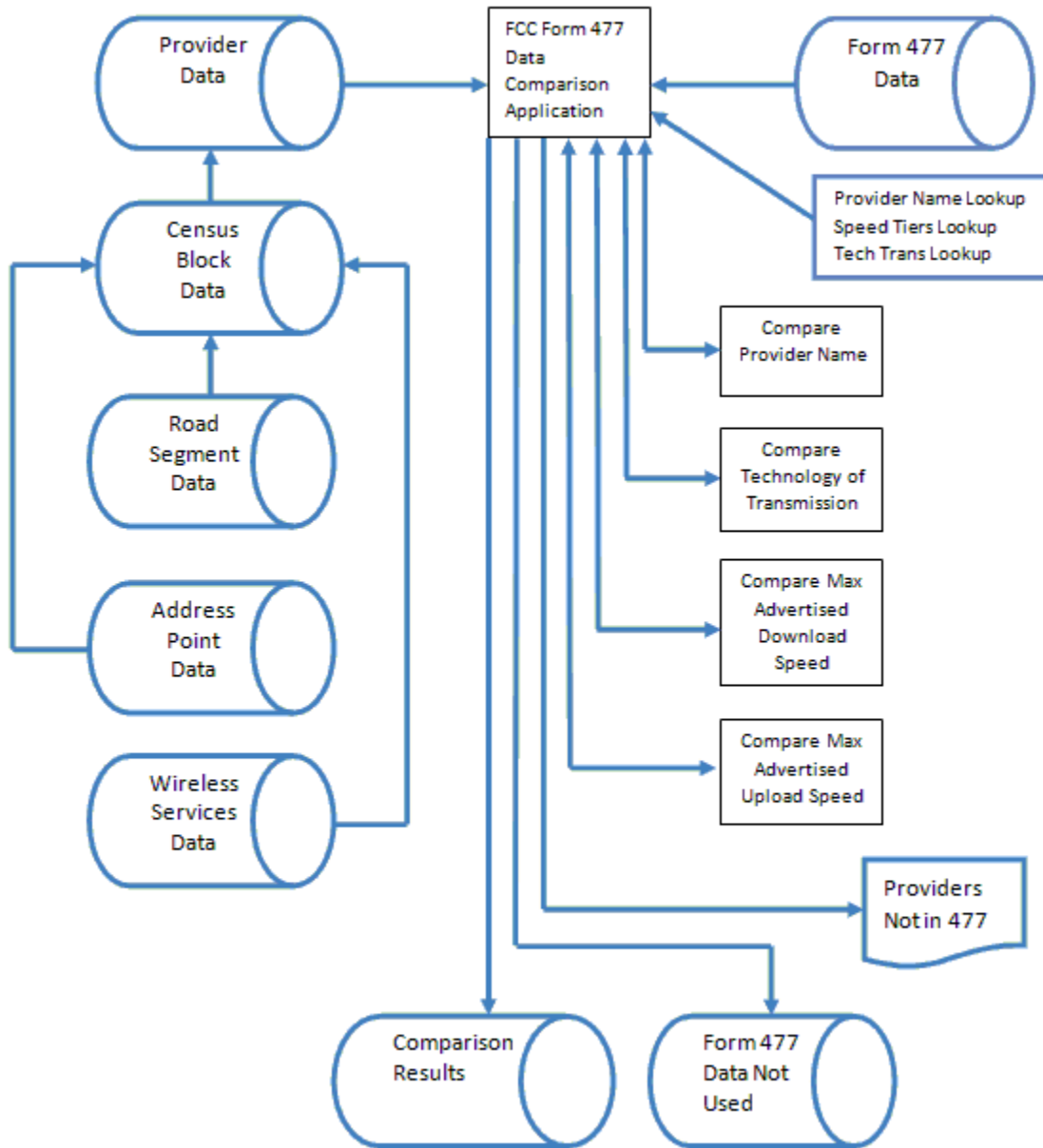| SBI Data	FCC Form 477
PROVNAME	PARENT_COMPANY
TRANSTECH	TECHNOLOGY_CODE
MAXADDOWN	DOWNLOAD_RATE_CODE
MAXADUP	UPLOAD_RATE_CODE

Process

Figure 2.   FCC Form 477 Comparison Process

FCC SPEED TEST

The FCC Speed Test dataset is built through the use of the Consumer Broadband Test tool found at http://www.broadband.gov/qualitytest. This dataset is updated prior to each SBI comparison assessment.

FCC Speed Test data is compared to the SBI Awardee-provided service census blocks, road segments, address points, and wireless service areas. This comparison source contains wireline and wireless data and compares provider names, technology of transmission, and speeds at the census block level. An IP address provider name lookup tool was created to verify that the correct Speed Test record is compared to the corresponding provider name. FCC Speed Test provider names, technology of transmission (wireline and wireless), and speed test results (KBPS to MBPS SBI Speed Tiers) differ from those of the SBI data, so lookup tables were created for matching purposes.

Comparison Source Preparation Description

The IP Address field for the FCC Speed Test data is put through an "IP to ISP" lookup routine to find the provider to whom the user subscribes. Any records that are not geocoded or have an intersected census block ID (vintage 2010) that does not match the first two digits of the State FIPS code where the point is located are ignored. Finally, the mobile and non-mobile records are loaded into the state comparison database.

Comparison Fields

Table 2. FCC Speed Test Comparison Fields

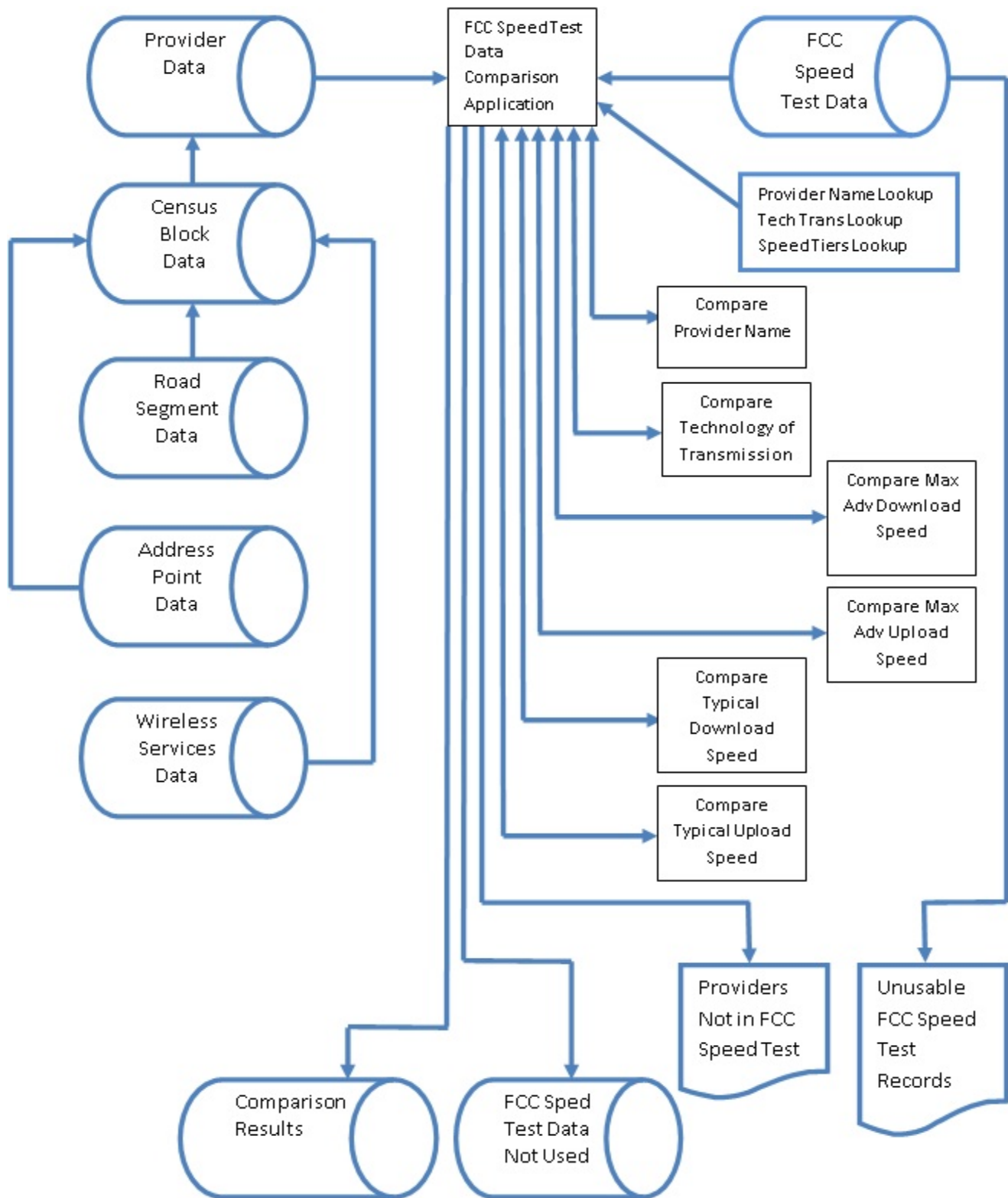| SBI Data	FCC Speed Test
PROVNAME	IPISP
TRANSTECH	CONNECTION
MAXADDOWN	DOWNLOAD_SPEED
MAXADUP	UPLOAD_SPEED
TYPICDOWN	DOWNLOAD_SPEED
TYPICUP	UPLOAD_SPEED

Process

Figure 3. FCC Speed Test Comparison Process

Mosaik Solutions

Mosaik Solutions sources its wireless coverage pattern information by working directly with the wireless operators as well as other publicly available sources of information.

Mosaik Solutions market research data is compared to the SBI Awardee-provided wireless service areas only. Mosaik Solutions provider names and technology of transmission codes differ from those of the SBI data, so lookup tables were created for matching purposes. The Mosaik Solutions data does not contain any speed information.

Comparison Source Preparation Description

The delivered shapefiles with protocols of EVDO, EVDO_REV_A, HSPA, UMTS, UMTS/HSPA, HSPA+, LTE, WiMAX, and WISP are loaded into a spatial database environment. The feature classes are then intersected with census blocks (vintage 2010) to allow for comparison to the SBI records. Duplicate records are removed by summarizing on market name, protocol, and census block. The Mosaik Solutions records are then loaded into the state comparison database.

Comparison Fields

Table 3. Mosaik Solutions Comparison Fields

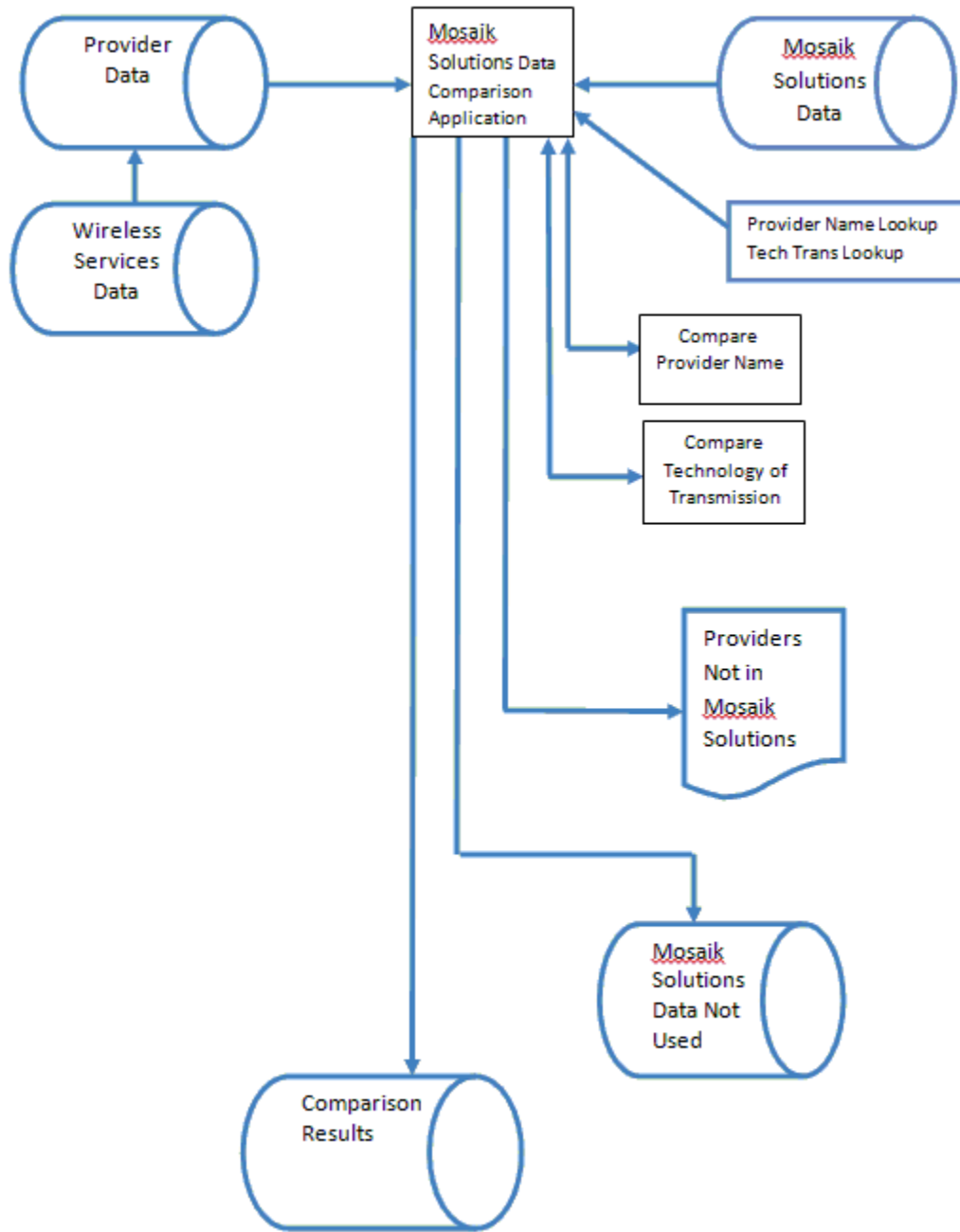| SBI Data	Mosaik Solutions
PROVNAME	MKTNAME
TRANSTECH	PROTOCOL

Process

Figure 4. Mosaik Solutions Comparison Process

Centris

Centris receives millions of household-level records from a provider of internet transactions where an address is associated with an IP address. These records were joined with Centris' Leading Telco, Primary Cable Provider, and Overbuilder Cable Provider Boundary files by block to identify data provider.

Centris market research data is compared to the SBI Awardee-provided service census blocks, road segments, and address points. This comparison source contains wireline data and compares provider names and technology of transmission at the census block level. Centris provider names differ from those of the SBI data, so lookup tables were created for matching purposes.

Comparison Source Preparation Description

Centris data is delivered in a table format containing a provider, technology, census block, and various other geo-location fields. The unique combination of provider, technology, and census block are loaded into the state comparison database.

Comparison Fields

Table 4. Centris Comparison Fields

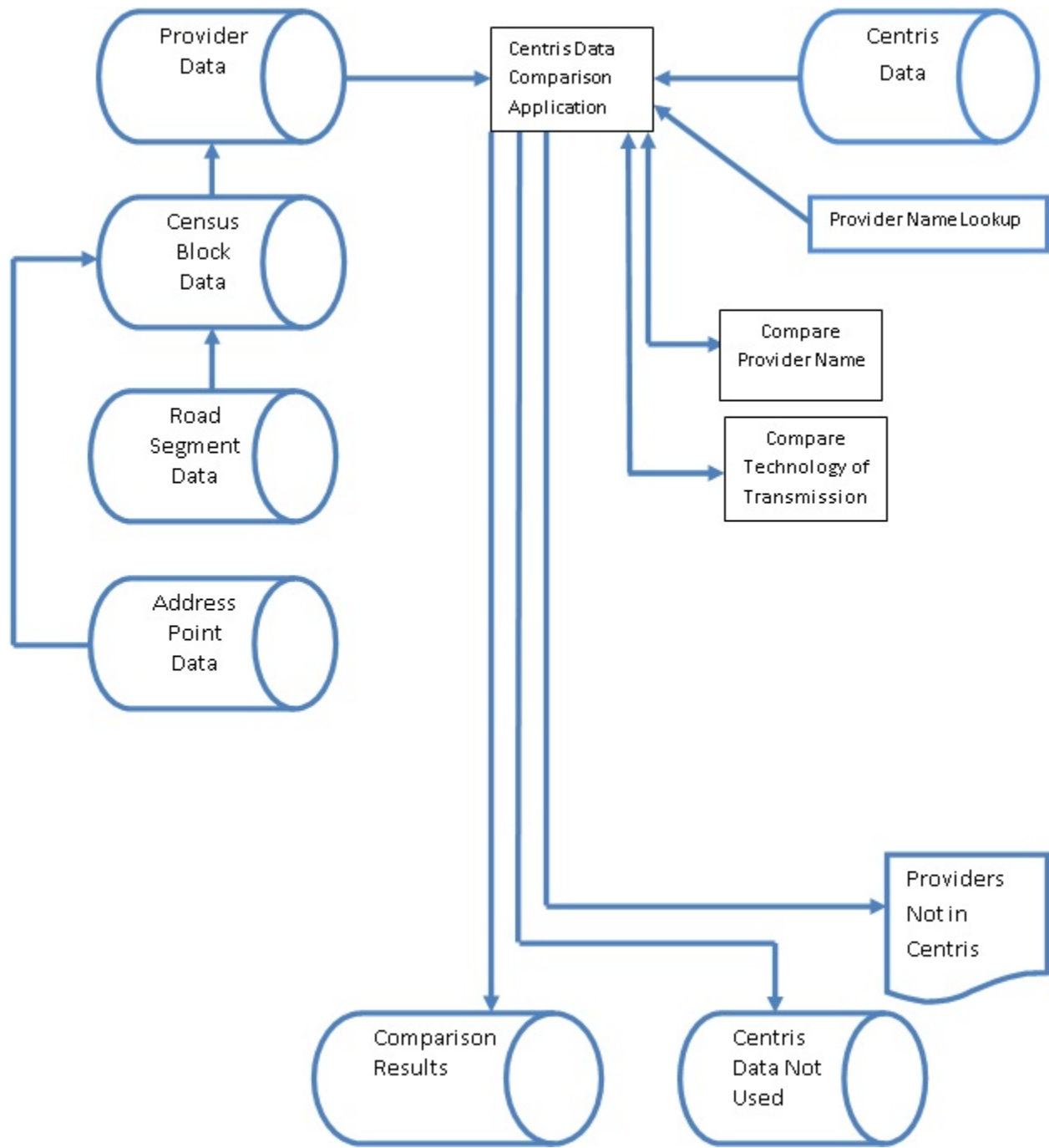| SBI Data	Centris
PROVNAME	PROVIDER
TRANSTECH	technology_of_transmission

Process

Figure 5.   Centris Comparison Process

Ookla

The Ookla dataset is a subset of Ookla's Netmetrics product which is built through the use of Ookla's SpeedTest.net tool that is licensed by thousands of organizations for primary testing of their LAN, WAN and VPN networks.

Ookla Speed Test data is compared to the SBI Awardee-provided service census blocks, road segments, address points, and wireless service areas. This comparison source contains wireline and wireless data and compares provider names, technology of transmission, and speeds at the census block level. Ookla uses Maxmind's provider name lookup tool so that the correct Speed Test provider record is compared to the corresponding SBI provider name. Ookla Speed Test provider names, technology of transmission (wireline and wireless), and speed test results (KBPS to MBPS SBI Speed Tiers) differ from those of the SBI data, so lookup tables were created for matching purposes.

Comparison Source Preparation Description

Ookla data is formatted in a manner similar to FCC Speed Test as they are both point features..   The only records used are those where the Mobile Device had GPS enabled while running the speed test. Each record then contains an accurate latitude and longitude that is joined spatially to its underlying Census Block. The database is then split to separate those records where WiFi was enabled to compare against wireline records. Finally, the Ookla records are loaded into the state comparison database.

Comparison Fields

Table 5. Ookla Comparison Fields

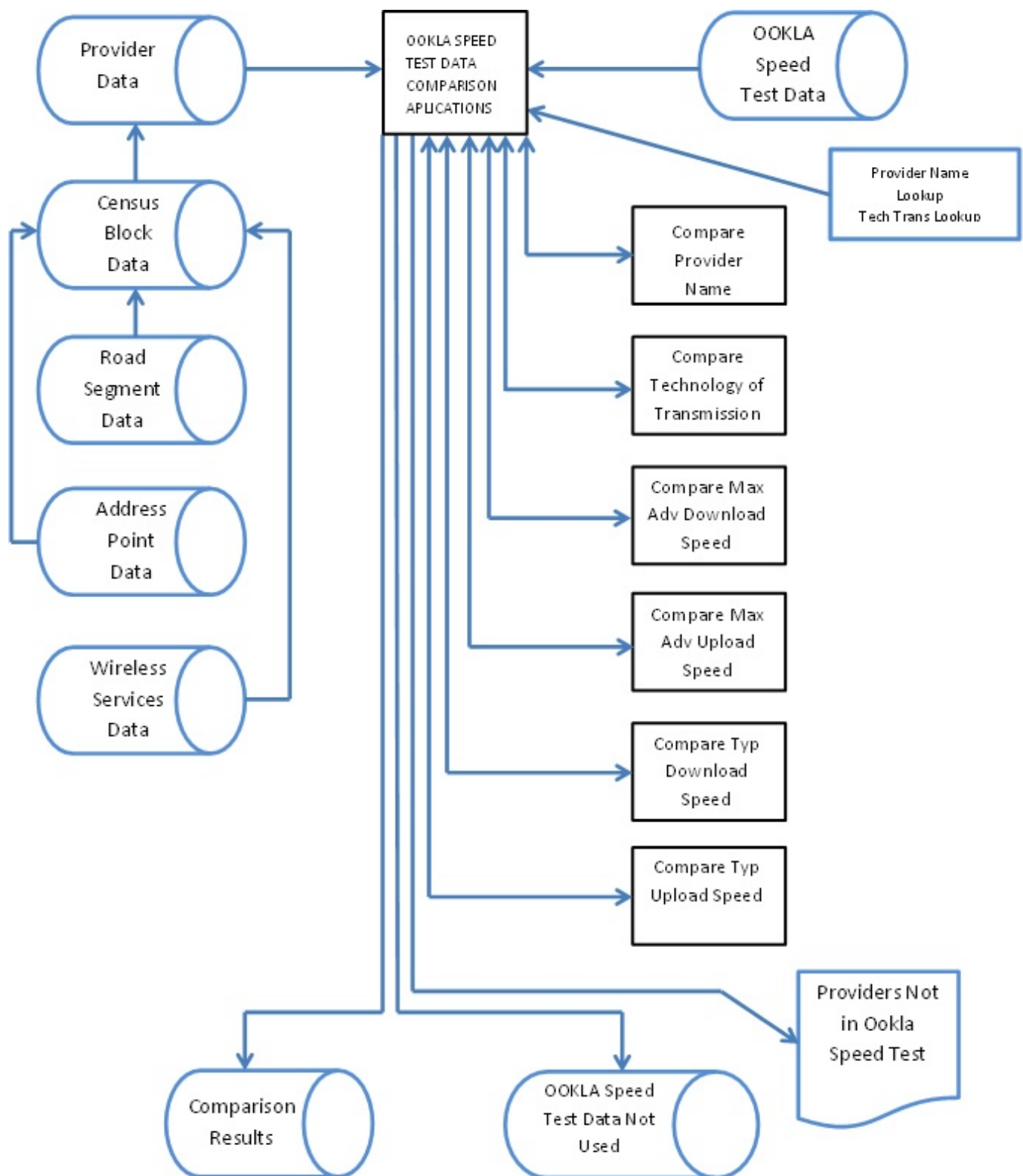| SBI Data	Ookla
PROVNAME	ISP_NAME
TRANSTECH	CONNECTION
MAXADDOWN	DOWNLOAD_KBPS
MAXADUP	UPLOAD_KBPS
TYPICDOWN	DOWNLOAD_KBPS
TYPICUP	UPLOAD_KBPS

Process

Figure 6. Ookla Comparison Process