CT Broadband Mapping

Data Processing Report

Supplement

Submission 3

March 31, 2011

CONNECTICUT PROGRAM OVERVIEW

In response to the Notice of Funds Availability published in the Federal Register on July 8, 2009 (NOFA), the State of Connecticut Department of Public Utility Control (CT DPUC) submitted a grant application for consideration under the National Telecommunications and Information Administration's (NTIA) State Broadband Data and Development Grant Program (SBDD), for broadband mapping. The CT DPUC, pursuant to Executive Order 32-A, has been designated as the single Connecticut state entity eligible to apply for funds under this program.

The State has long been committed to regarding broadband delivery and enhanced use as a fundamental goal. The State has developed a planning strategy to marshal the State's resources and stakeholders and establish Connecticut as a leader in broadband usage, in addition to being a leader in "e-Government" and other broadband-dependent endeavors.

The State entered its SBDD initiative not possessing any data related to broadband service, availability, or infrastructure that could readily support the requirements of the Broadband Data and Development grant program. Due to technical considerations, the DPUC has partnered with Applied Geographics Inc., to support the data collection and mapping efforts.

So far CT has been very successful in acquiring the requested information from the broadband service providers, and is utilizing this information on our own http://CT.gov/Broadband website as well as providing the needed information up to NTIA to support their national map.

SUBMISSION 3 OVERVIEW

For submission 3 (s3), roughly 50% of providers stated that their submission 2 (s2) service areas should be reused for s3. The other 50% gave modifications to their data, provided updates to their earlier submissions, or delivered entirely new datasets.

In general, the submission 3 processes followed the same basic approach that was used in submission 2. This document summarizes the following:

- Submission 3 Processing Assumptions
- Reference Data
- NTIA Submission Data Model Schema Changes

SUBMISSION 3 PROCESSING ASSUMPTIONS

Based on NTIA feedback and information provided in NTIA webinar sessions, the submission 3 data processing workflow is based on the following assumptions to meet NTIA submission requirements. Many of these are similar to s2.

1. All census blocks are mapped based on 2000 census blocks. Any data submitted in 2009 format was converted to 2000 for submission. During processing a 'hybrid' census dataset (2000 IDs with 2009 line work) was used to take advantage of the improved 2009 line work. Prior to

submission to NTIA, all features were mapped back to the 2000 census blocks. The Reference Data section below contains additional details.

2. For consistent representation the state road reference data used was 2009 Census Tiger Line IDs (TLIDs). Other data sources (non-TLID features, or 2000 TLID features) were mapped to 2009 TLID features.

3. Overview was removed completely from submission data due to the fact that all maximum advertised up/down speeds are being reported in blocks, roads, and wireless features. IN addition, none of the providers were willing to submit detailed pricing information.

4. Due to our NDA restrictions, address points and last mile points will not be submitted to NTIA.

5. Some providers did not submit middle mile elevation. Wherever possible, we went back to providers to obtain their middle mile elevation information.

6. Terrestrial Mobile Wireless and Terrestrial Fixed Wireless (licensed and unlicensed) were treated as wireless coverage and were delivered as a shape. In cases where a provider served the same technology and spectrum with different speeds, overlapping areas were removed and the higher speed was assigned.

7. The submission 3 Provider data model is currently based on the NTIA data model as of 1/13/11.

This section describes the reference data used in submission 3.

BLOCK REFERENCE SETUP

For s3, a hybrid block dataset (2000 IDs with 2009 line work) was used to take advantage of the improved 2009 geometry. The data was set up as follows:

- 2009 BlockID suffix is dropped and the blocks are dissolved (by Block ID) to produce data with 2000 BlockIDs and 2009 shape geometry
- Block size (AREA) is calculated combining the 2000 land area (ALAND) and water area (AWATER)
- AREA is converted from square meters to square miles to calculate square mileage (SMI).
- If the SMI of a block is less than or equal to 2, then the less than or equal to 2 square mile indicator (LE2SMI) is set to true.

ROAD REFERENCE SETUP

To take advantage of the 2009 geometry improvements, 2009 Tiger Line IDs (TLID) were used for data processing in s3.   Any non-2009 TLID (i.e. 2000 TLID or other) submitted by providers were mapped to the 2009 reference data. The data was set up as follows:

- The GT2SMI (Greater Than 2 Square Mile) indicator is set to True when:
    - The 2009 road segment <u>is completely within</u> a <u>hybrid</u> block that is NOT less than 2 square miles
- Only minimum and maximum address ranges and a single zip code for each road segment is maintained.

REFERENCE DATA SOURCES

The following summarizes block and road reference data sources for submission 3:

BLOCK REFERENCE DATA: 2009 CENSUS BLOCKS

The 2009 Census Block data is the most recent geometry provided by the US Census Bureau and has these characteristics:

- The full 2009 Block ID is made up of the following characters:
    - (2) State
    - (3) County
    - (6) Tract
    - (4) Block
    - (1) Suffix - The 2009 Census Block data allocates a one-character alphabetic suffix to the end of a 2000 Block ID for all blocks that have been subdivided.
- Fields of interest include:
    - [BLKIDFP]:: char(17) – Full Block ID
    - [ALAND] :: double(14) – Land Area
    - [AWATER] :: double(14) – Water Area
- The 2009 Census block geometry has been adjusted to correspond with the revised and amended 2009 Census road data.

- This data was downloaded for each state from the following website:
  http://www2.census.gov/cgi-bin/shapefiles2009/national-files

ROAD REFERENCE DATA: 2009 CENSUS TIGER LINES

The 2009 Census Tiger Line data contains the most recent geometry provided by the Census Bureau. The following is a list of characteristics:

- The Tiger Line Identification (TLID) system is stored as a double data type, although it contains only integer values
- Fields of interest include:
  - [TLID] :: double(10) -- (Tiger Line ID)
  - [FULLNAME] :: char(100) – (Full Name)
  - [LFROMADD] :: char(12) – (Left From Address)
  - [LTOADD] :: char(12) – (Left To Address)
  - [RFROMADD] :: char(12) – (Right From Address)
  - [RTOADD] :: char(12) – (Right To Address)
  - [ZIPL] :: char(5) – (Zip Left)
  - [ZIPR] :: char(5) – (Zip Right)
  - [ROADFLG] :: char(1) – (Road Flag – Is segment a road?)
- The 2009 Census Tiger Line road segment geometry was adjusted to correct 2000 segments misalignment; street name, type and directional information were concatenated into one database column (FULLNAME) and new road segments were added.
- The Census road data is packaged by county. Roads that exist as the boundary between counties will be duplicated in both county files.
- This data was downloaded by county as full tiger line data at the following website:
  http://www2.census.gov/cgi-bin/shapefiles2009/national-files
  - Source data was filtered by row were [ROADFLG] = yes to create the reference data set.

SUBMISSION 3: NTIA DATA MODEL SCHEMA CHANGES

The data model released on January 13, 2011 contained the following changes to the s2 data model:

- A new field was added to several feature classes called **Provider Type**
  - Provider Type is "Short Integer" and has domain values of 1, 2, or 3 (1= Broadband Provider, 2=Reseller, 3=other)
  - Most providers are calculated to be "1" (Broadband Provider). In some cases (e.g. State of Washington Public Utility Districts), providers are considered "Other" (value = 3).

- In the CAI feature class, the field **BBService** has been modified:
  - In S2 if the information was not known, the field was left blank (null)
  - In S3, if we do not have the information, Null's must be changed to code U (for Unknown) – nulls will not be allowed.

- Three new fields have been added to the CAI feature class. Wherever possible, these values have been populated in the CAI data.
  - **PublicWifi** (Y, N or U)
  - **URL**
  - **CAIID**

Due to Connecticut's geography and population, over 99.5% of the census blocks in the state are less than two square miles. The need for us to break apart coverage based on blocks versus roads leads to a lot of unnecessary confusion as well as creates some distorted pictures when you try to visualize this information on a map. For this reason, all of the maps available on the CT.gov/broadband website are published after we convert all of the data to just use blocks.

In the documentation form NTIA there has been a lot of discussion about making sure that a provider uses the same DBA and FRN consistently across all feature classes. We mentioned this to the providers, but there was some push back, especially from the national carries. In many of these cases the parent organization owns all of the middle mile infrastructure and utilizes one FRN for that. The actual distribution may often times be a smaller division of the organization, or may have been an acquired operation, and so the actual service will be listed under a different DBA and/or FRN.

There is also one provider that has given us two very similar DBA names: "DSL.net, Inc." and "DSLnet Communications, LLC". Each has a unique FRN. When we asked if we could merge these under one name, the company came back and was very adamant that we must keep these as separate and distinct entities. We honored that request and the data was submitted to NTIA as provided to us.

In regards to the NTIA sub domains concerning technology and speed, Connecticut has a few exceptions that we should note:

- Cablevision has reported to us that they serve an area with both Cable-DOCSIS 3.0 and Cable-Other, and that in this region "they advertise a maximum download speed of 101 MBPS (Tier 10)". We did ask the provider to clarify if the Tier 10 speed was for either technology type or only the DOCSIS 3.0 records? We also asked if there was an "Other" technology that might also deliver 101MBPS. Every time we asked, the provider simply responded with their original answer "they advertise a maximum download speed of 101 MBPS." So you will note that Cablevision has records with Technology codes of 40 and 41, with download speed tier 10.
- Verizon New York Inc. reported to us that they provide both fiber to the end user and ADSL service to some areas of the state. Most of their service is in fact fiber to the end user, but in the other locations they report that there is fiber run most of the way, but not into the actual premises. They are advertising this alternative as ADSL, with max advertised download speeds in tier 9.

On the NTIA webinar on March 17th, it was recommended that the states generalize their wireless data submissions, to include filling or dropping small areas, and reducing the number of vertices. Unfortunately due to the extremely late timing of this notice, we were not able to act on these recommendations. Many of our providers require in their NDA's that we process the data and then give them at least 10 business days to review and comment on any changes. Instead, we will look into adopting these changes in our fall submission.