Oklahoma Broadband Mapping

Data Submission Methodology Report

October 1, 2012

Oklahoma Broadband Mapping
_____

Data Submission Report (October 1, 2012)

Table of Contents

1 Introduction

This report is submitted along with the sixth data submission for the Oklahoma Broadband Mapping Project. This submission includes all data collected so far per the requirements of the National Telecommunications and Information Administration (NTIA) State Broadband Data and Development Grant Program (Docket No. 0660-ZA29) Notice of Funds Availability (NOFA) and formal and informal clarifications to it. Specifically, it includes broadband data collected from broadband providers and community anchor institutions data compiled from various sources for the State of OK. The State of OK has retained a mapping contractor, The Sanborn Map Company to perform the work related to the Mapping Grant for this project. Data from the previous submission is now publicly accessible via the OK Broadband Program (http://broadbandmapping.ok.gov/).

This document is a supplement to the five previous reports submitted with previous data submissions on May 1, 2010, October 1, 2010, April 1, 2011, October 1, 2011, and April 1, 2012 respectively. Therefore, it builds on the documents provided with those submissions. Rather than repeat the contents of the previous report, this document makes incremental updates on various topics where changes have been made in the methodology or reiterates the methodology used. Please refer to the previous documents for further details.

2   Overall Project Status

2.1   DATA COLLECTION

This section details data collection related to NTIA deliverables which include broadband data and community anchor institution data.

2.1.1 Broadband Data

For this submission, Sanborn started data collection efforts on July 5th, 2012 by sending out data update requests and technical data specifications. These were sent to a large list of companies which were compiled from multiple lists (FCC 477 list (dated June 30, 2011), Wireless Internet Service Providers Association (WISPA)) and from any providers that were identified through other sources such as web research, planning meetings, State outreach, etc. Sanborn also uploaded the final data for each provider in NTIA format from the previous submission on the Sanborn Provider Portal. The providers were encouraged to use the provider portal and update their information on it.

We followed the same contact and follow-up protocols as the previous submissions. In brief, this involved following up with already participating providers after sending them a letter requesting data updates. For newly identified providers, we contacted them three additional times and offered any/all support to make this as easy as possible. We provided a due date for submission but worked with providers who needed more time. If providers did not submit updated data and did not respond to our efforts to contact them, we reused their existing data.

The following are some of the important changes or no changes:

1. We continued to request all providers to provide us their speed information in mbps rather than as a speed tier. We did this in order to better validate the data, analyze served/underserved, and identify the breakdowns in speeds within a given tier. For this submission, 47% of the participating providers in OK have given us their speed in mbps rather than speed tier.

2. Like the previous submission, we also requested fixed wireless providers to provide us appropriate information to do propagation analysis. We had more success with this in the current submission. For those WISP providers that provided us the data to accomplish propagation, we used Radio Mobile to do propagation analysis and iterated with the providers until the parameters were suitably selected to get appropriate output. Propagation analysis results were provided to the providers for review through our provider portal and google kmz file formats to ensure validation. In OK, we were able conduct propagation analysis for six providers (AirLink, Cowboy.net, Omega 1 Wireless, PTCI, Plainsnet, and ProValue.Net) and collect propagation results from Link Technology for

one provider (HTS Wireless). Two providers submitted propagation results conduct themselves (Ruralinet, and Valnet).

3. We continue to not collect data from resellers.

4. Due to our NDA restrictions, last mile infrastructure points, if submitted by providers, are not being submitted to NTIA.

5. We continue to submit data for satellites in this submission based on NTIA clarifications. Four satellite providers have been identified in Oklahoma which are; Hughes, Starband, Wildblue, and Stratos. We have received data from all the companies listed above except for Stratos. We also were able to add an additional satellite provider, identified by NTIA (Skycaster) in this submission as well.

6. Due to NDA restrictions, address points are not included in this submission to NTIA for any commercial provider.

7. Some providers did not submit middle mile elevation or backhaul capacity, particularly when they asked us to reuse previous submission data. Wherever possible, we went back to providers to obtain that information, but it is not available for every record. In this submission, we also improved the elevation of middle mile data substantially. Some providers were providing elevation above sea level – we were able to work with the largest of these providers to resolve the issues and will continue to work on improving this over the next submission.

8. If a cable based wireline provider provides both DOCSIS 2.0 and DOCSIS 3.0 service to the same area, the block or road was listed only once with a technology code of 40.

9. Providers were only willing to indicate on a general level if they served business, residential or both, so we did not get any providers that broke down the type of service by block. Only if the provider stated they only serve business to business customers did we fill in the "category of end user" with a code of 2, otherwise this field was left blank. There are three providers in OK who are identified as serving business customers only. These are:

    1) Cogent Communications, Inc.
    2) TW Telecom of Oklahoma LLC
    3) XO Communications, LLC

10. This submission is being made based on the NTIA data model as of 08/09/12 provided by NTIA on the SBDD site.

    We have added nine new providers in this submission:

1) Cowboy.Net (fixed wireless)
2) Cross Valliant Cellular (mobile wireless)
3) Cross Wireless (mobile wireless)
4) Diamond Net (fiber)
5) HTS Wireless (fixed wireless)
6) Omega 1 Wireless (fixed wireless)
7) ProValue (fixed wireless)
8) Skycasters (satellite)
9) Wavelinx (fixed wireless)

In this submission:

1) We have contacted a total of 193 providers in OK, of which 55 providers were contacted for the first time.
2) We have identified 104 potential providers, of which 90 are participating in this map to date and 14 have refused to participate. In addition, 24 providers have not responded to our efforts to contact them and we are not sure whether any of these providers are actual providers or not. A list of the non-responders, resellers and non-providers is provided at the end of the document and all of these potential broadband providers were contacted. Even if some providers were identified as non-providers or resellers in previous submissions, we continue sending out data request letters to these providers in case their status has changed in any way.
3) 44% of the providers submitted new or updated data whereas for the remaining providers, we reused data from their previous submissions. This is in contrast to 49% providers submitting new or updated data during the previous submission.
4) We do not report areas of service for providers that have refused to participate or have not responded to our requests for data. In some cases program office staff are aware of approximate service areas for non-participating providers, but to date we have reported only areas that meet our validation criteria. If estimated service areas are desired we would collaborate with other states and NTIA to develop and disclose a workable methodology.

During this submission period, we had the following changes in providers:

1) Oklahoma Telephone and Telegraph changed its name to Oklatel Communications, Inc.
2) Wildblue changed its name to Viasat
3) Community Cablevision Company changed its name to Community Cable & Broadband

2.1.2 Community Anchor Institutions Data

The community anchor institutions data continues to be crowd-sourced through the online data gathering application created by the Sanborn Team. Following discussions with NTIA on ways to improve the CAI IDs via email and guidance on what datasets to use for IDs, we have incorporated IDs for libraries and schools where possible. The numbers of community anchor institutions that have responded so far is provided below:

Category	Name	Total in Submission 6	Total with Broadband Information in Submission 6
1	School - K through 12	1966	375
2	Library	211	175
3	Medical/healthcare	459	162
4	Public Safety	1794	316
5	University, college, other post-secondary	79	22
6	Other community support - government	507	91
7	Other community support - nongovernmental	16	2

2.2 DATA PROCESSING

We started with the following base data:

Census Blocks:

For this submission, Census 2010 data was utilized. The data was set up as follows:

- Block size (AREA) is calculated combining the 2010 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 Segments:

2010 Tiger Line IDs (TLID) were used for data processing for this submission. The data was set up as follows:

- The GT2SMI (Greater Than 2 Square Mile) indicator is set to True when:
    - The 2010 road segment <u>is completely within</u> a 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.

All data received went through the following processing steps:

1. **Triage:** All new data were quickly reviewed to understand what was received, and in what format. We also made sure we had all the required components for NTIA's data model, such as their FRN and advertised speed information. We also screened for any known issues that we might have seen before (such as Excel 2003 spreadsheets that cut off at 32k row).
2. **Ingest:** At this time the data is actually brought into our systems. Each provider is set up with a unique file geodatabase to store their information. Record counts of what was received are logged so that we can validate that we did not drop anything in processing.
3. **Data Processing:** In this step, the data goes through a number of ETL routines to convert the raw proprietary information into a format similar to the NTIA format. The exact routine utilized depends on how the data is received.
    1) When a wireline provider submits a service boundary, we select all the blocks and roads inside that shape.

2) If a wireline provider submits a customer address list, the points are geocoded, and then the appropriate block or road segment is selected.

3) If a wireline provider submits block and road information using Census data, we just make sure everything is formatted to the appropriate specifications.

4) If the wireline provider submits any type of road or line data that does not directly correlate to the TIGER data set, we convert the lines to TIGER by selecting the road centroid and spatially selecting the closest segment in our data set. If the road is in a block less than 2 square miles, then the block is selected. Some manual cleanup is also applied to make sure we do not accidentally drop any road segments that should have been processed.

5) Wireless provider data is formatted to ensure that there are no overlapping polygons with the technology type and spectrum. In addition the data is cropped to the state boundary.

6) After each round of processing, we make sure that we only keep unique records. A unique record is defined as having a unique combination of FRN, Block/Road ID, and technology type. If there are multiple records with different speeds, but all else is equal, than we select the maximum of the advertised speeds.

4. **QC Review:** All data are then sent to a different analyst to perform a thorough quality control review on the processed data set. Record counts are compared to what was submitted. The QC staff also makes sure the ETL scripts and routines populated all of the right fields.

5. **QA Review:** Data is then sent to another team for Quality Assurance Review. In this step the data is not only double checked against what was originally submitted, but it is also brought up inside standardized ArcMap templates that allow us to make sure our results make sense. This often involves comparing the new data set with prior submissions, as well as looking for any possible technology or speed anomalies and verifying against third-party datasets (as discussed in more details in the next section).

6. **Provider Review:** Processed data is all posted to a customized web-mapping tool we commonly refer to as the Provider Portal. All providers were notified once their data was available on the site, and were given five business days to review the data and respond. In this site, providers can log on and visually see their processed data in a map format. It also allows them to overlay their raw data to help them validate that we did indeed process things correctly. The provider portal also has a suite of markup tools that will allow the providers to edit their data, including adding or removing service areas, and making changes to the data attributes.

7. **Comment Processing:** All comments and feedback received from the provider portal is then reviewed and applied to the processed data set. This updated data set goes back through our QA and QC processes, and if time allows, back out to the Provider Portal, for the provider to review and sign off.

8. **Data Append:** After all of the individual data sets are processed and approved, we run an append process which merges all of the individual provider data sets into one geodatabase. This is also the point where our team will do any final transformations to get our working data model into the latest NTIA publishing format.

9. **Submission Comparison Check:** Starting with this submission an additional check was added to our quality review process. An application was written that compares the individual provider's unique data that is stored in their unique file to that which is stored in our final appended file and the NTIA submission data. Any variation in each of these data files in thoroughly investigated and resolved. This was done to assure no data loss or data transformation issues. We also compare the submission 5 dataset to the submission 6 dataset, review any variations and assure that the changes found can be documented as being requested by the provider.

10. **Final QA/QC:** A series of quality checks are run on the final appended data sets to ensure it is ready for submission to NTIA. We also run the latest version of the NTIA receipt tool at this time. If any issues are flagged as failing they are reviewed and corrected. All warnings are also reviewed and either corrected or documented in the attached document which explains that we have validated this data and any last issues are corrected.

11. **Submission to NTIA.**

2.2.1 Submission 6: NTIA Submission Data Model Schema Changes

The latest data model released was released on August 8, 2012 was very similar to the previous data model. No substantive changes were noted and changes related to allowable speed and technology of transmission combinations. Most of these combinations have exceptions to them and hence were not being completely disallowed by NTIA.

2.3 DATA VALIDATION

Sanborn has continued to perform the same validation on the data as the previous four submissions (details in previous reports and a summarized version provided below). Some minor updates to the validation process are discussed below.

1) QC of the data at various steps – this includes when data is received (triage), when it is processed through the various processing steps discussed above, etc.

2) Spatial checks against public and commercial datasets
   a. For OK, we continued to use the following datasets for validation:
      i. Exchange Boundaries: for DSL boundaries
      ii. MediaPrints: for Cable and Fiber boundaries
   b. We did not use speedtest.net speed data that we used previously for validation as we had our own speed test data that was more current and pertinent.
3) Speedtest data and other data collection for verification
   a. We continue to use speedtest data collected through our interactive map and community anchor data crowd-sourced for validation purposes.
   b. For this submission, we added an additional dataset to check against – FCC speed test data. We geocoded the data, used the IP to reverse engineer the provider name and used it to check speeds where possible.
   c. We also incorporated any feedback we received through the interactive map – this included feedback such as incorrect speeds, incorrect boundaries, missing provider or areas of no service, etc.
4) Verification by providers – processed data are uploaded on our Provider Portal for providers to review both the outcome of data processing and any issues that we found in the third-party and crowd-sourced validation. Issues pertaining to a particular provider are highlighted and shown in the portal for those providers only. Issues that are global and cannot be assigned to a particular provider are shown to all providers (e.g. there are no providers in this area, or we tried to get service here and heard x from A provider, y from B provider, etc.). Previously, we were highlighting these issues through a letter but in this submission, we have integrated the feedback through the Provided Portal. We make additional calls to providers who have issues.
5) Planning workshops and local validation –
   a. During this submission, local validation was undertaken by an independent group, the Center for Spatial Analysis at the University of Oklahoma (OU). OU performed an independent survey gathering data points from CAI's and the GIS community for the State of Oklahoma. Within Sanborn's validation process, OU's points were compared against provider's data. Those data points found in question were taken back to the providers for correction. Also, during this submission, Sanborn created an efficient way to make sure the CAI data that OU had gathered from phone calls or public gatherings was entered straight into the database without duplicating efforts for either team by designing a CAI data entry interface. This allows OU to increase their efforts to gather more data points by streamlining the process and provides current data to Sanborn to use for each submission and report to NTIA.

b. We have reviewed any issues that the State Planning team has identified and brought to our attention.

2.4   UNIVERSE OF CONTACTED PROVIDERS/NON-PROVIDERS

We have identified 104 potential providers, of which 90 are participating in this map to date and 14 have refused to participate. In addition, 24 providers have not responded to our efforts to contact them and we are not sure whether any of these providers are actual providers or not. A list of the non-responders, resellers and non-providers is provided at the end of the document and all of these potential broadband providers were contacted. Even if some providers were identified as non-providers or resellers in previous submissions, we continue sending out data request letters to these providers in case their status has changed in any way.

2.4.1 Non-providers
4D Networks Corp.
Atlas Telephone Company
Charter Communications
Comcast Cable Communications, Inc.
Cyber Rover
Fulltel
IO-2 Services
KoehlerPro Wireless
LightEdge Solutions Inc.
Magic Wireless Internet Service Providers LLC
McLeodUSA Telecom Services Inc. / PaeTec Corp
OKC Broadband (Ideal Advertising Inc.)
Oklahoma 5 Licensee Co., LLC
PCS Internet Services
Qwest Communications Company, LLC
Reach Broadband
Stouffer Communications / Granby Telephone
Telovations, Inc.
Texhoma Wireless
The Internet Shop
Tulsa MetroNet
United Wireless Communications, Inc.
University Corporation for Advanced Internet
UnplugUSA
Verizon Business Global LLC dba Verizon Business
Zayo Enterprise Networks, LLC

2.4.2 Resellers
Broadview Networks Holding Inc.
BullsEye Telecom, Inc.
Earthlink
Enventis Telecom Inc. / Hickory Tech Corp

Global Crossing Telecommunications Inc.
Logix Communications, LP
Metropolitan Telecommunications of Oklahoma, Inc.
New Edge Network, Inc.
Reallinx, Inc.
Telefonica USA, Inc.
TulsaConnect
Westel, Inc

2.4.3 Non-Responders/Difficulty Contacting

Airosurf Communications
Alliance Comm Network
Cable West
Coalgate Internet
CSWEB.NET
DataFlys
Datz
eConnect
Flash-Link Internet Service
Greenfly Networks, Inc.
HDR Internet Services/ OnALot.com
INETmax
KPowerNet, LLC/KAMO
Lakeview Cable
MEDIACOM LLC
ms bit
OneLink Wireless
Onlineok.com
ruralOK
The Junction
upperspace.net
Utopian Wireless Corporation
VectorLink
Wireless Broadband of Oklahoma

2.4.4 Not-Participating

Atlas Broadband
BartNET
EasyTEL Communications
eVolve Business Solutions LLC/Cincinnati Bell Inc.
horizon net
LRC Group
Meriplex Communications, Ltd.
OneNet
PriceNET Wireless
Rhino Communications
Stratos Offshore Services Company
Summit Digital, Inc.
Vroom Wireless, LLC

WEHCO Video, Inc.