Montana Broadband Mapping
Deliverable Report

Submitted To:

Chad Hultin
Project Coordinator
Montana Broadband Mapping
Montana Department of Commerce

Submitted By:

Fred Gifford
Tetra Tech EC Inc.
and
Ken Wall
GeoData Services Inc.

April 15, 2010

Contents

Executive Summary

The following report describes methods and issues related to the April 15th, 2010 deliverables to NTIA for Broadband Mapping in Montana. This data submission is meant to demonstrate our ability to produce data in the format required by NTIA and to provide NTIA with data that has been collected and validated to date.

Montana has elected to provide deliverables in the NSGIC data model format for submission to NTIA. The NSGIC database schema remains a work in progress, in our view is more of an operational model, and does not strictly comply with NOFA and NTIA state broadband mapping criteria. A best effort was made to fine tune this deliverable to match NOFA\NTIA requirements while still taking advantage of the NSGIC model. We anticipate feedback and modifications of this model before our final deliverable on June 24.

Montana has developed a very robust operational data model, components of which are described in this report, to support our broadband mapping efforts. We feel our operational model can support any reasonable modifications to NTIA requirements or the NSGIC data model. Since this deliverable format is derived from our operational data model, we anticipate some modifications will be required. We can take best practices recommendations from the NTIA April 21 webinar for instance, and incorporate those into the final deliverable without major modifications of our work flow and operating rules.

Our mapping process starts with infrastructure points (central offices, remote terminals, wireless towers and antenna locations, middle mile and backhaul), cable franchise areas, and anchor institution addresses. When providers have not supplied detailed information of their service areas that can be mapped at the census block level, coverage models are derived dynamically from this infrastructure based on geoprocessing techniques specific to each broadband technology. Examples of geoprocessing techniques include developing propagation models using the Longley-Rice model for wireless coverage and using infrastructure points in conjunction with the road network to predict the area served for DSL coverage.

We have developed a system to quantify "validated" data for the purpose of determining what is suitable for delivery to NTIA. The operational data model maintains reliability and validity codes, together with completeness checks to track which data elements are complete or still in process of refinement. Infrastructure is compared to public data, independent measurements, and telecommunications provider submittals at varying levels of geography. As more data is obtained from providers and systematically checked against infrastructure points, the reliability and validity progress from 1 (not validated or reliable) to 10 (validated and reliable). Completeness is primarily dependent on provider input, and can be supplemented in many instances with independent measurements. The process is iterative. Five of the providers included in this data set submitted infrastructure data at the address level. The remainder have submitted at a coarser geographic scale, most often by census tract, small scale paper or digital map, or generalized town location. Our validation methods provides the ability to use general

information and iteratively cross check and improve the coverage models as more accurate data is obtained.

This deliverable was intended to fully exercise the final NTIA deliverable submission by populating the tables and geographic deliverables for all validated data in a final format. We chose data that was defined in our operational schema with both reliability and validity values greater than or equal to 7 (on a scale of 1-10) for this evaluation deliverable. This was possible for valid portions of 14 of the providers that have submitted data so far. We did not limit the deliverable to "complete" data, since none of the data so far in Montana from any provider is 100% complete for all data attributes. All data in this submission will improve in validity and completeness between now and the final deliverable. It is possible that additional coverage will be defined for some of the included providers for infrastructure points that did not yet meet the validity limits we used for this deliverable. Approximately 23% of the providers in Montana have submitted some form of data to date. All providers who have not yet provided data have been contacted by phone and/or sent an email packages that include a spreadsheet to facilitate future data submittals.

Note: During review of this data submittal in the NSGIC format we noticed a small number of polygons have been dropped from the "FC_Census_Block_lessthan_2sqmi" feature class. We were able to repair the missing geometry but the attributes for these polygons remain blank. This problem does not exist in our operational data which is also stored as a file geodatabase. The majority of dropped polygons are very small and we are assuming the problem is due to a spatial extent\tollerance\resolution issue in the target geodatabase but have not verified that to date.

Reliability, Validity and Completeness

Reliability codes apply to the source data points and polygons and assess the authority of the source we obtained the data from and the level or coarseness of the geography (address or town). Validity codes are determined from cross checks of data sources and the number of independent sources of verification. These are as simple as comparing speed test locations against DSL modeled polygons, or as complex as geospatial analysis operations such as a kernel density function cluster analysis. Completeness is determined by public sources, independent measurements or provider submittals and checks on the domain classes required for the final NTIA deliverables such as Technology of Transmission domains, Speed Test domains and serving facility and wireless spectrum facility types and categories. The categories for these, and the subsequent records in our operational geodatabase tables grow and change as new data is obtained. We are maintaining these as feature level metadata tied to points and polygons maintained by analysts and technicians in a wiki table and coding them to the geodatabase. In this way the unique situations that arise can be cataloged and maintained with some level of flexibility while contributing to the final indices in a controlled fashion.

Reliability Codes

The two factors incorporated in reliability codes include the level of geography that was used as a source or provided as a clarification of location and the authority of the source for the information. We are also considering clusters of point information from independent measurements and sources to be higher in reliability than individual point information.

Generally, the coarser the source geography the lower the resultant score. Everything besides an address or street intersection, latitude/longitude location, or location provided in a georeferenced digital source is assigned a reliability score less than 5. This applies to source data coming (e.g. a central office located in a city instead of an address) and review comments on a previously mapped location (e.g. “That location is wrong, I know it is on the south side of town”).

We have incorporated the reliability code into our last point of aggregation (LPA) and provider coverage geodatabase files, and into some of the publicly available data (PAD) geodatabases. We are also carrying a short text field (50 characters) with a descriptive rationale for the score. This will allow us to focus more on the lower scores that need to be confirmed, and ignore the high confidence data scored as 9’s and 10’s.

Reliability Codes		
Code	**Description**	**Detailed Description**
0	Not assigned	• Not yet assigned
1	Level 1	• Checked but unverified
2	Level 2	• County • Presence by other coarse geography (e.g. administrative region)
3	Level 3	• City • Census tracts • Cable Plus (area likely to have been annexed into an incorporated town or CDP)
4	Level 4	• Cable - incorporated • Zipcodes • Census blocks
5	Level 5	• GeoTel unverified • Confirmed by provider or anchor institution key advisor but to geography coarser than address or intersection
6	Level 6	• Qwest/Midcontinent or other web site random testing check • Speed test from individual average residential
7	Level 7	• From anchor institution key advisor Webex • GeoTel verified address only with no 3rd party confirmation from public sources ○ Building unverified • Speed test from anchor institution

8	Level 8	• From provider
		• FCC ULS or ARS
		• Geotel verified address and possibly verified by 3rd party source (Google Streetview)
		  ○ Another provider's sign is on building (usually Qwest)
		• Geotel possibly verified by 3rd party source (NAIP, Google Streetview)
		• From state authoritative public data source (e.g. DCN or SummitNet)
		  ○ Address or building unverified
		• Speed test from cluster of average residential

9	Level 9	• From provider with authoritative confirmation
		• Geotel verified address and verified by 3rd party source (NAIP, Google Streetview)
		  ○ Provider sign on building
		  ○ Tower or dish visible
		• From provider or anchor institution check of our data * Root Wireless

10	Level 10	• From 2+ authoritative confirmations

Validity Codes

We include validity codes in the last point of aggregation infrastructure data which drives creation of the DSL models. We also include validity codes in each of the final technology of transmission deliverables for polygons and point feature classes. The scales of validity vary by each major type and function.

Infrastructure Validity Codes

The purpose of this validity code is twofold:

1. To determine which infrastructure points are turned into DSL model coverages
2. To use as a reference in other coverage validity checks

Infrastructure Validity Codes		
Code	**Description**	**Detailed Description**
0	Level 0	• Not yet assigned
1	Level 1	• Not yet assigned
2	Level 2	• Not yet assigned
3	Level 3	• Checked against MT PSC Report or DSLReports at the town level • Checked against SummitNet anchor institution data
4	Level 4	• Checked against two or more independent public sources at the town level • Checked against provider public data (e.g. Qwest ICONN) at the town level
5	Level 5	• Not yet assigned
6	Level 6	• Confirmation of DSL or cable from authoritative public data to broader geography than address not confirmed by provider
7	Level 7	• Authoritative public data at address level (e.g. Geotel) not confirmed by provider
8	Level 8	• Provider submission at the census tract level • Provider website independent address checks (Qwest, Verizon)
9	Level 9	• Provider submission at the census block level
10	Level 10	• Provider submission at the address level

Final Technology of Transmission Validity Codes

The purpose of this validity code is twofold:

1. To determine which elements are loaded in the spreadsheet provider submission packages in their review
2. To determine which provider coverages are chosen for submittal with one of the NTIA deliverables (April 15, June 24)

Final Technology of Transmission Validity Codes		
Code	**Description**	**Detailed Description**
0	Not assigned	• Not yet assigned
1	Level 1	• Unassigned at this time
2	Level 2	• Unassigned at this time

3	Level 3	• Checked against MT PSC Report or DSLReports at the town level • Checked against SummitNet anchor institution data
4	Level 4	• Checked against two or more independent public sources at the town level • Checked against provider public data (e.g. Qwest ICONN) at the town level
5	Level 5	• Confirmation of DSL or cable from authoritative public data
6	Level 6	• Provider website independent address checks (Qwest, Verizon) • Provider submission at the census tract level
7	Level 7	• Provider submission at the census block level • Provider submission at the census block level confirmed by Speed test cluster OR RootWireless
8	Level 8	• Provider submission at the address level
9	Level 9	• Provider submission at the address level confirmed by Speed test cluster OR RootWireless
10	Level 10	• Provider submission at the address level confirmed by Speed test cluster OR RootWireless

Providers in Submission

We have included 11 providers in this data check submission. Varying types and geographic levels of data were received from these providers. The table below shows the types of technology of transmission included in this deliverable.

Technology code	Description	Included in Deliverable
10	Asymmetric xDSL	Yes
20	Symmetric xDSL	No
30	Other Copper Wireline	No
40	Cable Modem--DOCSIS 3.0	No
41	Cable Modem—Other	Yes
50	Optical Carrier/Fiber to the End User	Yes
60	Satellite	No
70	Terrestrial Fixed Wireless--Unlicensed	Yes
71	Terrestrial Fixed Wireless--Licensed	No
80	Terrestrial Mobile Wireless	Yes
90	Electric Power Line	No, none expected in Montana
0	All Other (default)	No

A brief description of each provider and the type of data they submitted is provided below for context in understanding the deliverable -

3 Rivers Telephone Cooperative, Inc. (Technology of Transmission: ADSL, SDSL, FTTP, Other)

1) Pdf map of central office locations and subscriber locations, with county boundaries, exchange boundaries, and census tracts.

2) Pdf map of network schematic
3) List of census tracts with number of customers by upload and download speed tier (assume advertised rates) and technology of transmission.

Blackfoot Telephone Cooperative, Inc. (Technology of Transmission: ADSL)

1) Central office address
2) List of census tracts with number of customers by upload and download speed tier (assume advertised rates). Technology of Transmission is listed in the cover letter as DSL, but download speed tiers range up to 15 mbps and may include cable.
3) PDF map of census blocks in service area, central offices, and customer location points. Because the pdf file was locked we were unable to view map elements individually. The customer location points obscure the details of the census blocks.

CenturyLink, Inc. (Technology of Transmission: ADSL)

1) List of LERG switch locations (wire centers) with VH coordinates and addresses.
2) List of census blocks with Technology of Transmission.
3) No information on speed tiers was provided.
4) No information on the number of subscribers was provided. In her cover letter Mary Taylor said "This request is inconsistent with the NOFA clarification and is considered competitively sensitive. Therefore no information is being provided."

Lincoln Telephone Company, Inc. (Technology of Transmission: ADSL)

1) Central office address
2) List of census tracts with number of customers by upload and download speed tier (assume advertised rates). Technology of Transmission was included in the cover letter.

Mid-Rivers Telephone Cooperative, Inc. (Technology of Transmission: Cable, DSL, FTTP, Fixed Wireless)

Cable & Communications Corporation

1) Central offices and wireless towers in two DWG files.
2) List of census tracts with number of customers and Technology of Transmission.
3) No information on speed tiers was provided.

Northern Telephone Cooperative, Inc. (Technology of Transmission: ADSL)

1) List of central offices and DLC sites with addresses or lat/long coordinates
2) List of census tracts with number of customers by upload and download speed tier (assume advertised rates). Technology of Transmission was included in the cover letter.

Qwest Corporation (Technology of Transmission: ADSL)

1) Provided a polygon coverage of their coverage with each polygon indicating the speed tier. The speed tiers appeared to be vastly overstating the speeds we have observed so far from speed tests.
2) Provided large circles representing their central office infrastructure, which when converted to centroid locations appeared to be slightly dithered locations of their actual offices, with a 1000-2000 ft random offset at a random azimuth. They were close enough to the verified points we had collected from public sources to allow us to do a provider verification of infrastructure.

Range Telephone Cooperative, Inc. (Technology of Transmission: ADSL, SDSL)

1) Pdf map of central office locations with telephone exchange areas (approx scale 1:1,200,000)

2) List of census tracts with number of customers by upload and download speed tier (assume advertised rates). Technology of Transmissions were included in the cover letter but were not identified by census tract. Typical upload and download speeds were also provided in the cover letter, but only for the entire service area.

Southern Montana Telephone Company (Technology of Transmission: ADSL)

1) Central office city location
2) List of census tracts with number of customers by upload and download speed tier (assume advertised rates). Technology of Transmission was included in the cover letter.

Triangle Telephone Cooperative Association, Inc. (Technology of Transmission: DSL, FTTH)
Central Montana Communications, Inc. (Technology of Transmission: DSL, FTTH)
Triangle Communications (Wireless)

1) List of central offices with lat/long coordinates
2) List of census tracts with number of customers and Technology of Transmission.
3) No information on speed tiers was provided.

Verizon Wireless (Technology of Transmission: ADSL, Wireless)

1) No information from provider. Towers from FCC database used, and identified as Verizon through preliminary examination of Root Wireless independent measurements. Preliminary examination of this data was also used to populate the typical speed attributes.

NSGIC Data Model

The April 5, 2010 version of the NSGIC data model was used for the format of this deliverable. Some modification of domain tables were made to be compatible with NTIA requirements. These tables are derivatives from our operational data model, which is much more extensive and includes many additional attributes that are not required in the final NTIA deliverables. Whenever we had incomplete or unknown data we left the value null. One of our recommendations is to assign missing values rather than null to these attributes.

The NSGIC data model differs from the schema descriptions in the Federal Registry. We anticipate the NTIA/FCC will provide an authoritative version for use for the final project deliverables. Review comments reflect the 4/5/2010 posting of the NSGIC data model. Our comments on the NSGIC data model primarily apply to how the data models differed from the Federal Registry specifications for deliverables. The NSGIC data model isn't wrong, it just didn't reflect all the requirements as stated in the Federal Registry. The NSGIC data model appears to be more of an operational data model that supports the development of deliverables rather the format specified for the deliverables. The Utah revision states that it is intended to satisfy the NITA deliverable requirement, and that the user may need to define additional operational tables and feature classes.

Observations and Comments on NSGIC Domains

- Domain names need to be consistent with NOFA and the authoritative source.
- The "Serving Facility Backhaul Type Code" domain is utilized in two tables (Last mile and middle-mile/backhaul) in the Federal Registry. Perhaps the domain should be generalized to a "Serving Facility Type Code" domain and not be table specific.
- The "Serving Facility Codes" domain is defined differently in two tables in the Federal Registry. To clarify this in the data model perhaps the domain should be specific to the table in this situation. "Serving Facility Codes Last Mile" and "Serving Facility Codes Middle Mile".
- The NSGIC "Spectrum Used" domain is not a valid domain in the Federal Registry.
- There should be a Y/N domain as this is a restricted domain in multiple fields in multiple tables. For example, in the Community Anchor Institutions table the field broadband service should be Y/N domain.
- The Record Format for Residential Broadband Service Pricing and Speed Characteristics by County for Each Provider - recommend the State FIPS and County FIPS codes for each State's implementation should be a domain.

Observations and Comments on NSGIC Tables and Feature Classes

If the NSGIC geodatabase is intended to characterize NTIA/FCC deliverables then the object type must be compatible with the authoritative source in differentiating when a deliverable is intended to be a table and when it is intended to be a feature class.

- For example, Availability by Service Address is identified as a table not a feature class (no geometry, provided per State or Territory).
- The NSGIC Wireless Services not Provided to a Specific Address schema does not match schema in Federal Registry.
- NSGIC Census Block less than 2 sq. mi. and Record Format Road Segment Data are operational format (feature class) but deliverable is table. We kept these in feature class format. Utah NSGIC discussions mentioned not clipping or intersecting roads to census blocks, for this deliverable we did intersect the road segments.
- Feature classes that are identified in the Federal Registry as deliverable in a shapefile format need to have field names specified that are compatible with the restrictive field name domain of shapefiles (length and characters allowed).
- What is the role of the Sequential ID field in record format for Address Data for Each Provider if no end-user addresses are being collected?

Observations and Comments on Comparison of the April 8 Utah data model against the NSGIC posting from the first week of April, and what has been derived from the Federal Registry.

- The Utah and NSGIC geodatabase both differ from the Federal Registry descriptions. Most of the differences are minor in scope, but may impact data management procedures. The one exception is that Utah and NSGIC both include the end-user fields in Table 1a that were deleted. Didn't see any documentation as to why they have chosen to retain those fields.
- The Utah model redefines the data type for the FRN field from an integer to a text field to handle leading zeros in FRN values. Don't know if this was an oversight on the part of NTIA, but probably need to retain leading zeros on FRNs. The NSGIC data type for FRN is still a long integer.
- The Utah model creates a new feature class called "Service Overview" that combines some components of Table 1a, and Table 2a. This did not exist in the NSGIC model. The new feature class adds two new geography fields to track the geography represented. This is a significant change from the Federal Registry and the NSGIC data model.
- Utah redefined the Number Served in Table 1b to be number of potential end-users rather than current subscribers, which changed the NSGIC data type from a short integer to long integer for this field.
- Utah adds an "ID" field to all feature classes to hold a feature-level unique persistent identifier for a given state. There is a note that a state ID field will likely be needed to ensure uniqueness of records nationally. Probably need to add a “State” field to tables to ensure uniqueness nationally.
- For Table 1b (wireless) in the Utah model the spectrum field is now a domain and the descriptions define the type of service, which makes more sense. The NGSIC model did not define a domain.

- Utah did add a "Y/N" domain to be used by the Anchor Institutions table, as well as new domains to support geography type.
- Domains and fields need to be reviewed to address null values and unknowns. If a field is required, but other information in the record is complete, is it assumed that the incomplete records will not be included in deliverables? The data model should explicitly identify which fields are required, and which fields can be delivered as incomplete.
- There is a comment in the Utah model that says Census roads are NOT to be split for reporting and that sides of roads do not need to be differentiated.

General comments on Utah data model:

1. The Utah model specifically identifies all "BB_" feature classes as "deliverables" and leaves it to the user to define additional operational tables. Therefore, the Utah data model only contains what it considers to be deliverables.
2. The Utah model has good documentation and metadata that can be leveraged.
3. The Utah model does a good job of documenting unresolved issues by feature class, which need further review.
4. The Utah model defines field names that appear to be consistent with shapefile exports, but did not thoroughly review.