NOMINATION SUMMARY

Beelines: A Quality Control Methodology for Addresses

ABSTRACT OF THE PROGRAM

In 2017, the Montgomery County Department of Technology Services (DTS) Geographic Information Systems (GIS) team mapped lines (Beelines) connecting building footprints and building address points that fall on street centerlines. DTS-GIS maintains buildings for public safety E-911 CAD (computer aided dispatching), that includes updating addresses, building footprint shape (approximately 330k), 10 fields of pertinent attributes and attaching available PDF files (showing building access points, staircases, etc.). Planning such maintenance requires delineating for staff where building addresses fall on the street centerline as well as the distance from that point on the centerline to the actual location of the building footprint. Such tasks can be accomplished programmatically using desktop GIS software to create Beelines between points of duplicate addresses. By viewing beelines, possible errors in addresses can be ascertained by analyzing why the beeline is extremely long, crossing a centerline or crossing another Beeline.

DTS-GIS compiled several databases to develop and analyze address data for the purpose of data integrity. The databases included building footprints, building address points, centerlines with address ranges, geocoded points from building addresses, aerial photography, and beelines. To accomplish producing quality addresses for databases, such as centerlines and building footprints, beelines were used to highlight possible address errors.

THE PROBLEM OR NEED FOR THE PROGRAM

To enable the provisioning of quality, timely and accurate Public Safety services, DTS-GIS must provide centerlines with address ranges for routing purposes. To attain correct network routing lines, “fishbones” were created between the building footprint and where it landed on the centerline to determine if there needed to be an assist for network routing. In doing this project we uncovered tools to benefit an accurate centerline database for addressing, building address points for addressing and address anomalies.

DESCRIPTION OF THE PROGRAM

When DTS-GIS staff first encountered the idea of errors in using the extension, Network Analysis, Loudon County had created a data set of “fishbones” to remedy the problem. The DTS-GIS team members discussed the problems incurred by Public Safety’s software. When inputting an address, the possibility of a structure too far off from its fronting street and too close to a non-fronting street would result in routing errors. It was agreed that including “extra” lines into the centerline database would assist in routing, showing access to buildings that had traits of difficult locations when routing. This idea created the notion of centerline assist lines, although this came with its own set of challenges for maintenance of centerline data. Difficulties, such as

trying to replicate a routing error and then producing assist lines programmatically, were the first hurdles to overcome.

It was determined to proceed with discovering a way to highlight addresses/buildings that may route incorrectly. To carry out this idea, the term “fishbones” was used to discover where these addresses existed. The idea was to connect the polygon buildings to the fronting street with lines. It was not clear however, which buildings needed these “fishbones,” so we decided on giving all buildings “fishbones” and worked our way backwards by flagging the “fishbones” that were not longer than a specified length. This methodology became a tedious task and in trying to uncover a more efficient way to proceed, a new reason to create ”fishbones” emerged and we changed the term from “fishbones” to “beelines.”

The discovery of “beelines” and its alternative use from the original idea of assisting centerlines in the routing process, emerged from the geographic patterns “fishbones” were creating. Examples of patterns being displayed were “beelines” crossing other “beelines,” “beelines” an extremely long length reaching across the extent of the County, and “beelines” crossing centerlines. The process in which the “beelines” were created should not have resulted in any of the patterns described above. In seeing these patterns that should not exist, the following project, which can be described as researching why a certain pattern was created by the “beeline” and what needs to be done to correct the pattern, made in error, developed. The result of the investigation into the patterns provided another method/tool in quality control and maintenance of building and centerline addresses.

Beelines were created and displayed using ESRI’s ArcMap 10.5 software and data sets, such as building addresses, building polygon/points and centerlines. Building addresses are maintained for Public Safety, are contained in a geodatabase (.gdb) file and are a culmination of many different addressing sources. Building polygons are used to show the location of a structure’s footprint based on aerial photography and building points are created from the polygon’s centroid, both containing address attributes. Centerlines are the premier addressing tool/data set for Public Safety routing, as it contains attributed address ranges on segmented lines for the purpose of finding a location in the County. Utilizing these data sets and the software, a new data set was created, called beelines, to highlight errors in addressing for quality control purposes.

The creation of the new data set, Beelines, was a process of combining building address points, and geocoded address points that locate on the centerlines and a line to connect these two points. The building address points are attributed with addresses, which were used to create the database (.dbf) file of standalone addresses. The centerlines that contain attributed address ranges that are on the left and right side were used to create address points on the centerline segments. In so doing, 2 sets of points were created: building locations and their corresponding address location

on the left and right side of the centerline. Having these data sets, building address points and address located points, ArcMap tools were used to create lines between them, in turn creating a Beeline.

When the Beelines were created, the lines displayed were expected to be linear and straight in fashion, indicating a direct path from the building point to the location of the offset address point on the centerline. On closer inspection, some beelines created patterns across the County extent, such as zig zag beelines, beelines crossing each other and beelines crossing centerlines. These patterns gave rise to questions, such as what caused the patterns, what type of patterns indicated a need for research and what patterns were an acceptable exception to the research.

The cause of the patterns that the Beelines made were a combination of duplicate addresses, geographic curve of the centerline and incorrect addresses. Duplicate addresses were a product of multiple building polygon/points that are on one piece of property, i.e. a detached garage or shed and a main house, creating a zig zag line. The geographic curves in the centerline may cause Beelines to cross the centerline because the beeline is straight and the address points are connecting to each other over the curved centerline. Incorrect addresses are the outfall to creating the Beelines and correcting the address so that the beeline does not crossover itself or centerlines. These 3 patterns that were relevant to this methodology of highlighting possible problems with addresses were now a stepping stone for initiating research.

Research began by viewing Beeline patterns, such as the zig zag line, on a macro scale to narrow down what patterns needed closer inspection. When viewing the County extent, a line zig zagging showed many duplicate building addresses connecting to one point on the centerline. These beelines were discarded as problems if the zig zag only displayed connections between building address points on one piece of property. The majority of zig zag lines were discarded as errors because it was determined if the addresses were duplicated, that was the reason for the ill formed line. Zig zag Beeline patterns were determined not to be an issue for incorrect addresses and in the attribute field, “Note,” it was marked as 0, to indicate nothing to be done about the building address or cline address. (See Figure 1)

Beelines that crossed each other were another possible indication of incorrect addresses. Most of these lines were displayed on cul-de-sacs, or centerlines with curved geometry. The line drawn from the building address point to the address on the cline point may cross other beelines, because they are straight lines reaching over each other to make the connection. The intentional beeline crossing another beeline were reviewed by checking the addresses from the building and the centerline points and were found not to be an error by address. These lines were then tagged in the attribute field, “Note,” as 0 to make this type of line to be ignored. (See Figure 2)

Beeline patterns that crossed multiple centerlines indicated five different scenarios. The first scenario begins by knowing that all addresses on a cul-de-sac are even or odd or consecutive number, and that this causes beelines to cross over the centerline to reach its intended address target and so this pattern is acceptable (See Figure 3). Scenario number two, a building at the end of a cul-de-sac, could be odd or even, but the position of the building and the curve of the centerline creates an intersecting situation; this scenario is also acceptable (See Figure 4). Scenario number three are the address ranges on the centerline that must be flipped (See Figure 5a and 5b) or the direction of the centerline needs to be flipped (See Figure 6), or the nodes of the centerline need to be adjusted to correct the address ranges (See Figure 7), a scenario needing to be flagged for resolution. The fourth scenario, where a building address is incorrect, in this case the components that make up a building address, i.e. premise number, street name or zip code is incorrect, this beeline is then lagged for correction (See Figure 8). A fifth and final scenario is discovery of exceptions to the address rules (for example, an odd address falls on the left side of the road or an even address falls on the right side of the road based on addressing not following the address rules (See Figure 9)). These beeline patterns are of most importance to highlight address errors, whether they be in a centerline address range or a building address, so that corrections can be made and a superior data set can be maintained for locational and networking projects.

USE OF TECHNOLOGY

The process was completed using ArcGIS desktop software, including ArcMap 10.5 and ArcCatalog 10.5. As mentioned above, ArcGIS' Arctools were used to simplify the data creation process, and the existing data sets were divided into the north and south County to allow the beelines to be edited by multiple users simultaneously. Two DTS-GIS staff were provided with reference data sets such as, property (MNCPPC, Maryland - National Capital Park & Planning Commission), aerial photography (DTS-GIS), Master Address Points (MNCPPC), Centerlines (DTS-GIS), Public Safety building polygons (DTS-GIS), and building points (DTS-GIS). Additional resources to check for assurance of addresses were United States Post Office, Google Street View and Zillow.com.

THE COST OF THE PROGRAM

DTS-GIS already possessed the necessary licenses for ArcGIS desktop software as well as internet access. A jurisdiction that needed to purchase such licenses to do similar work would have to pay $9,900 for the *ArcGIS for Desktop Advanced* license from ESRI.

The RESULTS/SUCCESS OF THE PROGRAM

The results of the beeline QC method effect the correction and maintenance of addressing data sets such as, centerlines, building and zip codes. Centerlines are the premier method in locating addresses for a variety of uses, primarily because the range of attributes assure that a wider number of addresses can be located. Buildings are a more specific way of locating addresses

because of the one to many relationships between address records. Zip codes are a dataset component of addresses and need only a smaller amount of attention in the correction/maintenance tasks, but are needed to break ties between duplicate addresses with the same street name. The success of the program has been a more accurate set of data for locating addresses in a process called geocoding which is used on many occasions to create, analyze and make decisions.

WORTHINESS OF AWARD

The program is award worthy because it provides a better address data set, makes it easier to catch and correct errors in address data sets and expedites the editing process. Several County departments and groups of people use address location for projects, and therefore it is vital to have accurate data. This is particularly true for the life-saving public safety E-911 CAD operations. This method of quality control highlights data that needs correcting. The editing process becomes quicker and takes staff less time to make corrections and changes. Beelines as a quality control method affords staff the confidence in providing accurate address data sets.

SUPPLEMENTAL MATERIALS (attached)