Drainage basin delineation from vector drainage networks

Craig Dillabaugh
Canada Centre for Remote Sensing
588 Booth Street
Ottawa, Ontario, Canada
craig.dillabaugh@ccrs.nrcan.gc.ca

ABSTRACT

An automated process for the delineation of drainage basins, using as input only the vector based drainage network, is described which provides an alternative to DEM based methods. The procedure involves an approximation of the line Voronoi diagram for the network using a constrained Delaunay triangulation of the network's Planar Straight Line Graph. The constrained triangulation, which includes all network channels as edges in its resulting triangles, is further sub-divided into a set of facets. A two-stage algorithm is then applied to these facets to assign each facet the identifier of a channel in the input network. The drainage area of each channel is delineated by merging all facets assigned to that channel to form a single polygon. The drainage basin for any point in the network can be determined by tracing the channels upstream from that point and merging the associated drainage polygons. Several examples of the procedure's application are provided on drainage networks of varying size and complexity. While the procedure cannot ultimately match the accuracy of a DEM based basin delineation, it can be used in cases where DEM data is either unavailable or of poor quality.

Keywords: Delaunay triangulation, basin extraction, drainage basins

1. Introduction

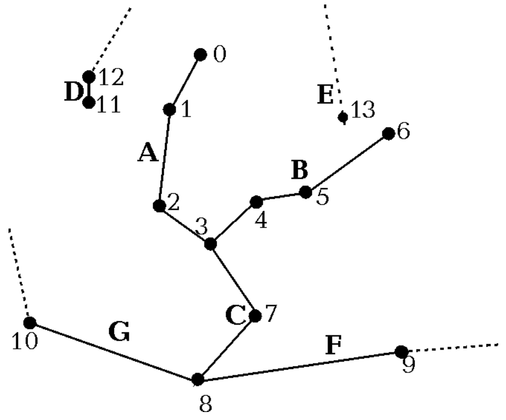
The extraction of drainage networks and their associated drainage basins (or catchment areas) is normally accomplished through the use of digital elevation models (DEMs). Whether the DEM is represented as a set of points in a Triangular Irregular Network (TIN), or as a raster surface, numerous techniques exist which allow the extraction of stream networks and their associated basins. Examples of such techniques can be found in Mower (1994) for raster data, and in Jones et al (1990), and Nelson et al (1994) for TIN based DEMs. However, if the vector representation of the stream network is available but the DEM is either unavailable or of inadequate quality then it may be desirable to estimate drainage basin boundaries based on the network vectors. Techniques for this procedure are not widely available. The purpose of this paper is to present a simple technique which allows drainage basin boundaries to be estimated from stream vectors.

In the absence of a DEM or any specialized models which might estimate basin shape based on network configuration, the most logical way to delineate the basins would be to assign to each stream segment the area which is closer to that segment than to any other. In other words, the distance between a point within the basin of a stream segment and that stream segment should be shorter than the distance between the point and any other stream segment. The technique presented here uses a partitioned Triangular Irregular Network (TIN) model, created from the input drainage network vectors rather than DEM data, to produce a set of polygons representative of the drainage area attributable to each stream segment in the network. Conceptually it is similar to producing a line voronoi of the stream network, though the actual polygons are not true line voronoi's but rather an approximation. A discussion of the production of true line voronoi diagrams can be found in Held (2001).

2. Conceptual Framework

2.1 Network Topology

The drainage network, which may include channels and lakes and perhaps islands (thought this was not tested) is initially stored in a representation which will be referred to here as an arc-node network. A small portion of such a network is shown in figure 1, along with its tabular representation. The individual stream channels (or lake shorelines) are represented as arcs (labelled A, B, C ...) which themselves are represented by a listing of vertices which mark the end points of a series of straight line segments which compose the arc. The vertices themselves are simply represented by their x and y coordinates. For the purposes of the discussion here, each vertex can also be considered to have a unique identifier (labelled 0 through 13 in figure 1). The start and end vertices of any arc are referred to as *nodes* and may be shared among multiple arcs (i.e. vertex 3 in figure 1).

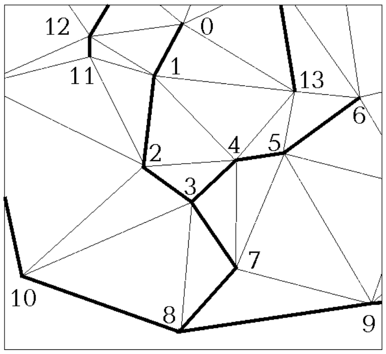


Arc	Vertices
A	0,1,2,3
B	3,4,5,6
C	3,7,8
D	11,12,....
E	13, ...
F	8,9, ...
G	8,10,...
....	....

Figure 1: Arc-node representation of network channels. Note that the network extends beyond the extent shown here, and individual arc segments extending beyond the display extent are indicated as dashed lines.

The procedure outlined here provides a means of drainage basin delineation by assigning *drainage areas* to each arc in the input arc-node network. The drainage area is a polygon which will represents the extent of the area in the physical landscape which drains to its associated channel. Actual drainage basins can be determined for the nodes in the network by searching up stream from that point to identify all channels that drain to it. The combined drainage areas for all those channels comprise the drainage basin for the selected point. The procedure of determining the drainage areas for individual channels is implemented by applying a constrained Delaunay triangulation to the input drainage network. This triangulation results in a set of triangles which are further subdivided into six *facets*. The drainage area for the channels was determined by assigning each facet to a channel, and then merging all facets given the same channel assignment into a single polygon. A Delaunay triangulation is a triangulation whose vertices are the set of input points and where no input point falls into the interior circumcircle of any other triangle in the triangulation. A constrained Delaunay triangulation is similar, but with the requirement that each segment in the input Planar Straight Line Graph (PSLG) serves as an edge in the triangulation. As such, a constrained triangulation is not a true Delaunay triangulation as the circumcircle requirement cannot be enforced (Shewchuk, 1996).

Each arc in the arc-node network can be broken down into a set of arc 'segments' which are straight lines connecting each adjacent pair of vertices in the arc. This collection of segments forms the PSLG for the arc-node network and serves as input to the triangulation procedure. The PSLG is represented by two lists, the first of these being a listing of all vertices in the network which consists of a unique vertex identifier and the x and y coordinates of each vertex. The second list forming the PSLG is a list of segments, where each segment is represented by the identifiers of two vertices that define its endpoints. Applying a constrained Delaunay triangulation to the PSLG creates a third list, which for each triangle simply records the identifiers of the three vertices of which it is composed. The vertex list for the input PSLG and the resulting triangulation are identical. This property of the constrained Delaunay triangulation, along with a second property, that being that each input segment in the PSLG forms an edge in the resulting triangulation, are essential for the proper functioning of the facet assignment procedure described below. In terms of the original arc-node network, these properties of the constrained Delaunay triangulation mean that each segment in the original arc network forms the edge of one or two triangles and that no triangle edges intersect any network arc. The constrained Delaunay triangulation of the network from figure 1, along with a partial listing of vertex identifiers for the triangles is shown in figure 2.



Triangle	Vertices
0	1, 2, 4
1	2, 3, 4
2	4, 5, 13
3	4, 5, 7
4	3, 7, 8
....	....

Figure 2: Constrained Delaunay triangulation for channel network along with its tabular representation. Segments from the original network are heavy lines, while the segments generated as part of the triangulation are light. Vertex identifiers correspond to those in figure 1.

One other feature of the triangulation is that it includes no information regarding the positioning of the vertices from within the original stream network. The PSLG records only the vertex positions and which vertices are paired to form line segments. The triangulation in turn only records which vertices compose each triangle. In order to enable the facet assignment procedure it is essential that information be recorded regarding the position of the network vertices in the original arc-node network. This information can be obtained at the time the PSLG is generated and is stored in what will be referred to as the *incident arc list*. The incident arc list records for each vertex in the vertex lists which arcs in the arc-node network that vertex belonged to. The incident arc list for the network in figures 1 and 2 is provided in table 1. The procedure for generating the incident arc works as follows; as each vertex is output the arc to which it belongs is recorded in the incident arc list for that vertex. A special node table is kept which records those vertices that are also nodes. When a node vertex is to be output, this table is checked to determine if that node has already been written to the vertex list (as the node of another arc). If the nodes has already been written then the current arc's identifier is added to that vertex's set of incident arcs without reinserting the vertex.

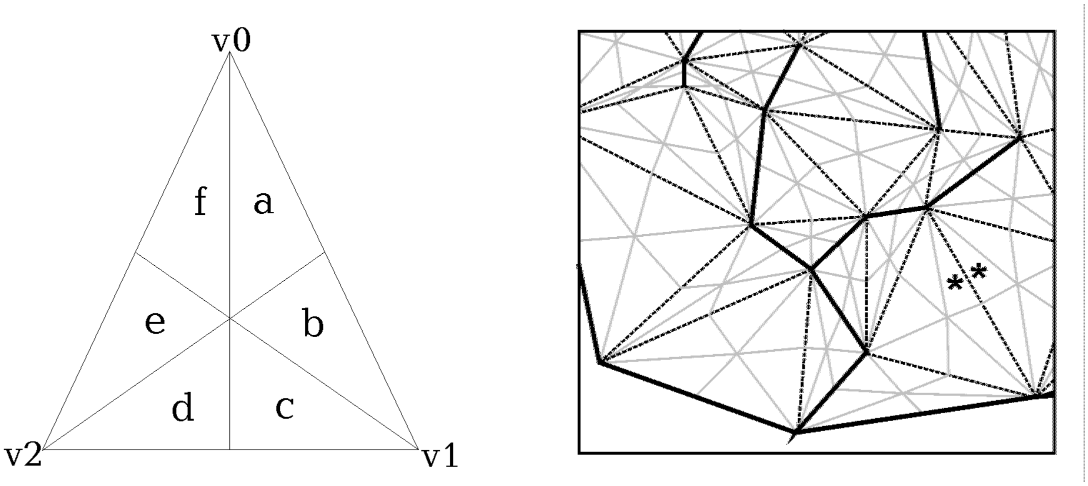
VERTEX	INCIDENT ARCS
0	A
1	A
2	A
3	A,B,C
4	B
5	B
6	B
7	C
8	C, G, F
9	G
....	.....

Table 1: Incident Arc List for Network shown in figures 1 and 2.

3. Basin Extraction Procedure

Following the triangulation step basin extraction was performed by further sub-dividing each triangle into six facets and assigning the individual facets to the arcs representing the stream segments or lake shorelines. The required inputs at this stage include the listing of vertex positions, the triangle file indicating the three vertices belonging to each triangle, and finally the listing of incident arcs. The facets are formed by simply identifying the triangle centre and dividing the triangle into six sub-triangles (or facets) by connecting lines from this centre point to the three vertices and between the mid-points on the three triangle edges as shown in figure 3a. The method of determining the centre point of the triangle is described hereafter.

In order to simply the description of facet assignment it is helpful to define a number of terms with respect to facets and their associations with each other. In figure 3a. facets ***a*** and ***f*** are considered to be 'adjacent' to the vertex v0, while ***b*** and ***c*** are adjacent to v1, and ***d*** and ***e*** are adjacent to v2. Furthermore facet ***a*** and ***b*** are considered to be adjacent to the line segment $L_{v0,v1}$ while ***c*** and ***d*** are adjacent to $L_{v1,v2}$ and ***e*** and ***f*** are adjacent to $L_{v2,v0}$. Within the triangle each facet has two neighbouring facets, its *adjacent neighbour* being the facet adjacent to the same vertex, and its *opposite neighbour* being the facet adjacent to the same line. Thus in Figure 3a facet ***b*** is the opposite neighbour of facet ***a*** while facet ***f*** is facet ***a***'s adjacent neighbour. Finally the triangle produced during the triangulation within which a facet occurs is referred to as that facet's *parent triangle*. The edges of each parent triangle can be shared with other parent triangles, in which case the two facets sharing both the edge and a vertex are termed *exterior neighbours*. This is demonstrated in Figure 3b where the two facets marked by the * are exterior neighbours.



(a) Single triangle with facet labels.

(b) Triangulation sub-divided into facets the stars indicate exterior neighbour facets.

Figure 3: Facet labelling.

Facet assignment is a two-stage procedure, with an initial assignment stage which is followed by a clean-up stage which iteratively assigns the remaining facets until all have been assigned. The value assigned to each facet is the identifier of an arc in the original arc-node network. Often features will occur in a network to which drainage should not be assigned, for example a coverage neatline, or small islands within a lake or river. Such features should be identified prior to processing and marked as invalid for assignment.

3.1 First Stage

Each facet is considered in turn, and is assigned to a particular arc (stream segment) according the following rules:

a. If the vertex adjacent to the facet has only a single incident arc (which is not invalid for assignment), then the facet is assigned to that arc.

b. If the vertex adjacent to the facet is incident only to a invalid arc then the facet assignment is set to 'reserved'.

c. If the adjacent vertex has multiple incident arcs then look at the line segment to which the facet is adjacent and;
  - *i. If that line segment is an arc in stream network then assign the facet to that arc number.*
  - *ii. If the line is not part of the stream network, then the facet assignment is set to reserved.*

Whether (a) or (b) applies can be easily determined from the incident arc list of the triangle vertices. For (a) there will be only a single valid incident arc number, and for (b) there will be no valid incident arcs. For (c) determining whether to apply (i) or (ii) requires the determination of which triangle edges (lines) belong to a particular arc in the stream network. Fortunately this information can be obtained from the incident arc list as a triangle edge represents part of the stream network when both vertices contain the same incident arc within their incident arcs list. Assuming clean topology of the original arc-node network no two vertices will ever share more than one common incident arc.

3.2 Second Stage

To complete the basin delineation all of the facets marked as reserved in the previous stage must still be assigned. The assignment of these facets is achieved by applying the following set of rules:

a. If no facets at all have been assigned in the current parent triangle then examine each facet's exterior neighbour in the adjacent parent triangles to determine if they have been assigned. If so give the current facet its exterior neighbour's assignment.

b. If there are assigned facets in the parent triangle then;

i. If the current facet is adjacent to a vertex with no valid values in its incident arc list then examine first the adjacent neighbour and then the opposite neighbour and when one of those neighbouring facets is found to have an assigned value, assign that to the current facet.

ii. If the current facet is adjacent to a vertex with valid incident arcs, then examine first the adjacent neighbour and then the exterior neighbour and if one of these is found to have an assigned value, assign that to the current facet.

Stage 2 is repeated until all facets have been assigned an arc number.

While the logic behind most of the rules outlined in stage one and two are straightforward some of the rules in particular stage 1 c.ii and stage 2 b.i and b.ii. can benefit by some further explanation. Figure 4 below is provided to demonstrate instances where these rules would be applied. In this diagram the facets labelled 1 through 4 are all adjacent to the vertex with an incident arc list which includes arcs A, B, and C. During stage one facets 1 and 4 are assigned to channels B and C respectively according to rule c.i, however, facets 2 and 3 cannot be assigned since the triangle edge they border on is part of the triangulation but not part of the network. Facets 2 and 3 are assigned during stage 2 according to rule b.ii. Facets 5 and 6 are adjacent to the junction of arcs A, B, and C, and will be assigned to arcs B and A respectively on the second iteration of stage 2, once again in accordance with rule b.ii. However, unlike facets 2 and 3 these facets will be assigned based on their exterior neighbours because the triangle to which they belong does not border on either of the arcs directly.

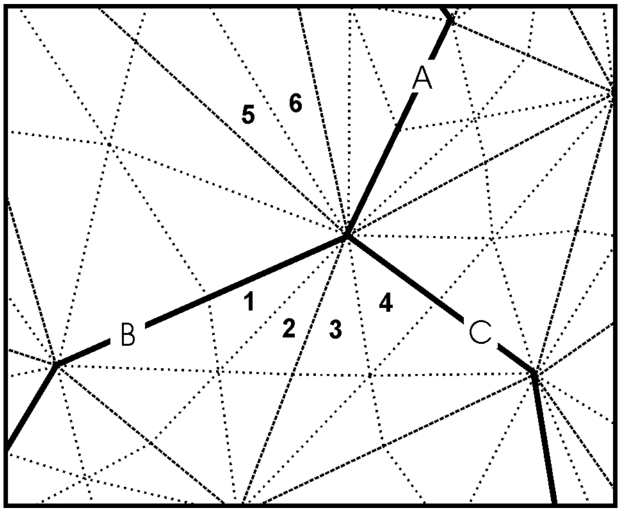
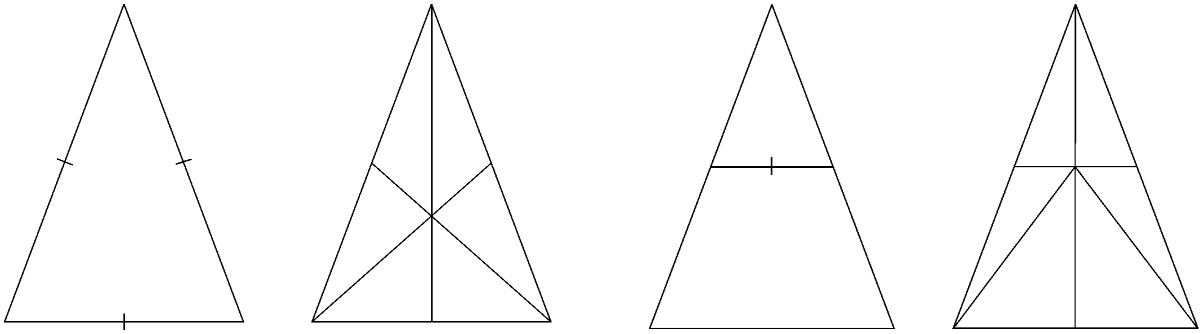


Figure 4: Solid lines represent original stream network, dashed lines represent triangle edges, and dotted lines represent facet edges. Channel (Arc) ID's are signified by uppercase letters while numbers identify the facets.

3.3 Determining the Triangle Centre-Point

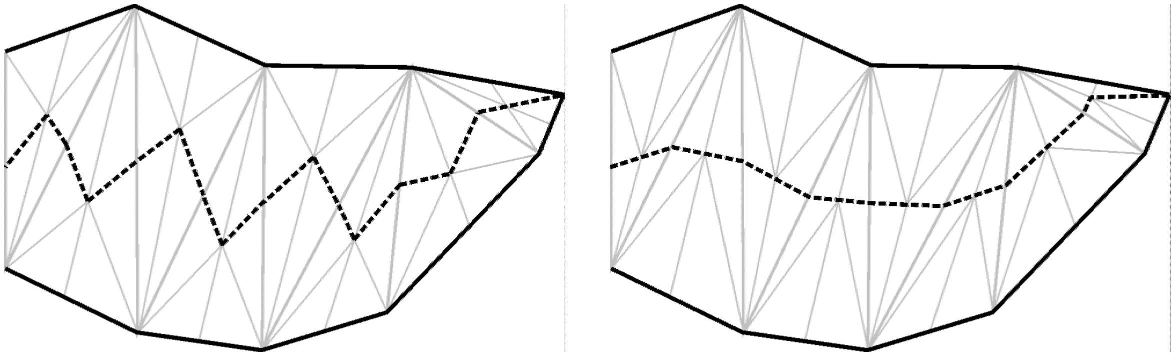
Division of the parent triangle into facets involves creating six line segments linking the parent triangle's centre point to its three corner vertices and the midpoints of the its edges. The first method for determining the triangle centre, which will be called the centroid method, is calculated by taking the arithmetic mean of the x and y coordinates of the three triangle vertices. A second technique for determining the triangle centre called the bisection method was also employed. The bisection method works by selecting two of the three triangle sides and making the triangle centre the midpoint of a line connecting the midpoints of these two sides. In general the method will work by simply selecting the two longest sides, however it is possible to use any two sides to determine a centre.



(a) Centroid method for triangle centre determination. (b) Bisection method for triangle centre determination.

Figure 5: Centroid and Bisection methods of determining triangle centres for facet generation.

For the most part the bisection method produces superior results, particularly in cases where stream segments run nearly parallel to each other. The reason for this is that when segments run parallel the triangulation produces numerous long thin triangles that alternate between having one and then two vertices on either of the stream segments. Because the centroid method takes the arithmetic mean of the vertices the centre is placed closer to the stream segment having two of the current triangle's vertices. Following facet assignment this produces the jagged division of triangles shown in figure 6a. Using the bisection method however produces a much smoother line (see figure 6b.).



(a) Centroid Method (b) Bisection Method

Figure 6: Division of facets between two stream segments using the centroid and bisection method. Solid black lines represent the original streams, grey lines represent triangle and facet boundaries, while the dashed black line represents the division between facets assigned to the lower and upper stream segments.

One area in which the centroid method tends to produce a better result is at stream junctions. This can be seen in Figure 6 where the centroid method actually produces a better division for the triangle occurring in the junction, while bisection does a better job for the remaining triangles. Thus rather than just selecting one method and using it on all triangles it is better to choose the centre selection method dynamically depending on the current triangles positioning relative to the stream network. The rules for determining which technique to used can be summarized as follows:

- If the triangle has two vertices on one stream segment and the third on another (as with the long narrow triangles in Figure 6), then bisection should be employed using the two lines originating from the lone vertex on the second stream segment.
- If all facets are the same then they will eventually be merged into a single drainage area so centre selection is irrelevant, however, the centroid computation is slightly simpler so use it.
- If the triangle has one vertex on one stream segment, one of the other, and the third vertex is shared, then this is a junction vertex so use the centroid technique.

3.4 Drainage Area Delineation

Once the facet assignment is complete then basin's are delineated by merging all facets given the same assignment into a single polygon. This merging can be accomplished using a vector based geographical information system. However, in order to speed the process of merging polygons, rather than output every single facet as a polygon, only the facet segments (edges) which were shared between facets having differing assignments were output. It is then left to the GIS to build polygon topology from the output facet segments. This technique further requires that for each possible assignment value a single label point be output which records the assignment value. A major assumption being made here is that ALL of the facets assigned to any one arc are contiguous, this of course will be the case so long as the drainage network is properly constructed such that no two different arcs have the same identifier.

4. Results and Discussion

For testing purposes five network coverages were selected ranging in both scale and database size. The smallest coverage in terms of area and size (number of arc segments) was the MACKENZIE coverage, which covered a small area along the Mackenzie Mountains which form the drainage divided between the Yukon and Mackenzie rivers. This network includes areas in both the Yukon Territory and the Northwest Territories in northern Canada. The next largest coverage was the BANKS coverage, which included the whole of Banks Island in the Northwest Territories. The largest coverage in terms of aerial extent was the North American drainage network NORTH_AM, which provided a small scale generalized drainage network for the entire continent. The final two coverages, and the largest in terms of database size, were the ALBANY and NELSON coverages, which included the drainage basins for the rivers emptying into Hudson Bay, Canada. Each coverage is named after its major river, but both include a number of small rivers which also drain into Hudson Bay. Figure 7 below shows two of these networks BANKS and NELSON.

(a) BANKS drainage network. *(b) NELSON drainage network.*

Figure 7: Drainage networks for BANKS and NELSON coverages.

The drainage network and the resulting basin polygon coverages we handled using the Arc/Info GIS. Arc/Info was used to ensure the quality of the input networks and to handle the polygon creation from the output facet boundaries and labels. Arc/Info also supports performing constrained Delaunay triangulation, however, the program TRIANGLE (Shewchuk, 1996) was used instead. The main advantage of TRIANGLE, which is freely available over the world wide web, is that the storage formats for its inputs and outputs are open and clearly documented and thus could be relatively easily integrated into the basin extraction procedure. TRIANGLE requires as inputs a simple PSLG and outputs a listing of triangle vertices . The facet assignment procedure was implemented in an external program written in the C++ programming language. All analysis was performed on a Sun Ultra 10 workstation, running the Solaris OS.

The processing involved first ensuring correct network topology for the drainage network and proper directionality of arc (stream) segments. Where proper directionality exists in the network it is possible to extract the complete drainage basin for any point in the network using the basin extraction procedure outlined hear. Following this stage the network was exported from Arc/Info and processed to generate the listing of incident arcs (stream segments) along with the PSLG for input into TRIANGLE. A small external, program, called ARCTRI, was written to achieve this. TRIANGLE was then run on the PSLG to generate a the constrained Delaunay triangulation, which along with the incident arcs file served as input to the facet assignment program. Results from the facet assignment were then imported to Arc/Info where polygon topology was constructed from facet outputs. The resulting polygons represented the drainage areas for each stream segment in the initial network coverage. The drainage basin for any point in the network could then be determined simply by tracing back through all arcs upstream of that point and selecting the associated drainage areas for the selected arcs. The processing times for triangulation and facet generation/assignment are shown for the five coverages in Table 2.

Network	Size (# Stream/Arc Segments)	Approximate Running Times	
		Triangulation	Facet Generation/Assignment
MACKENZIE	232	1 second	less than 1 second
BANKS	1,978	1 second	less than 1 second
NORTH_AM	6,426	9 seconds	3 seconds
ALBANY	45,714	32 seconds	3 seconds
NELSON	76,681	71 seconds	9 seconds

Table 2: Network Coverage Processing Times

Figures 8 thru 10 present the results of the extraction procedure for sub-basins in three of the network coverages. Figure 8 presents results for a delineated drainage basin from the ALBANY network. Just to the north (above) of this highlighted area the Attawapiskat River can be seen. One feature of interest in this case is the small area draining to a lake in the southwest end of the basin that is not marked as part of the basin. Since the lake is not connected directly to the drainage network its associated basin is not assigned to the basin as it likely should be in this case. A reasonable solution to this problem would that any when a drainage basin is delineated any drainage area polygons that are completely enclosed should automatically be assigned to that basin, however, this was not implemented here.

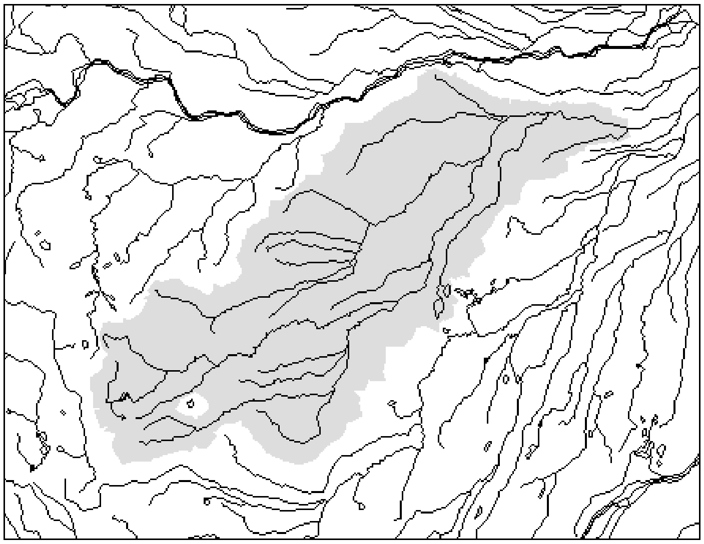
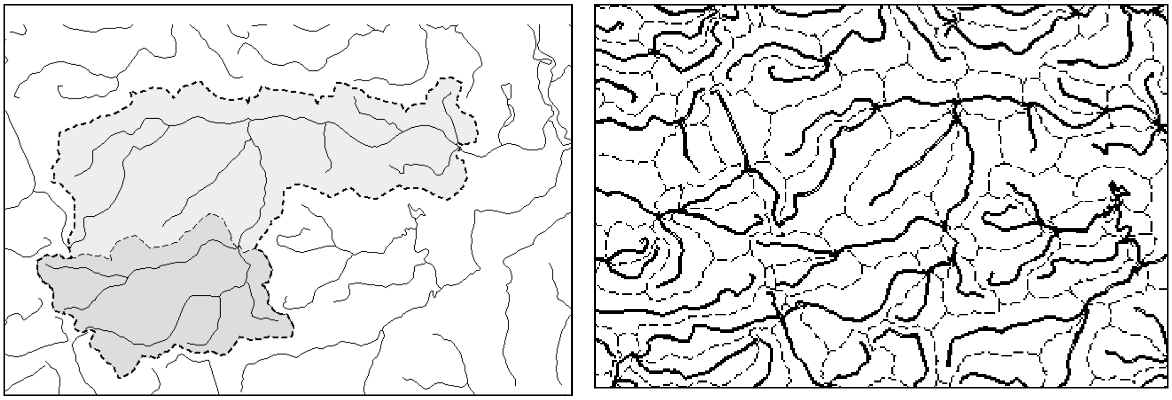


Figure 8: ALBANY network extracted drainage basin. The highlighted area flows to towards Hudson Bay to the east (right) side of the scene.

In Figure 9a a basin extracted from the MACKENZIE network is shown along with one of its sub-basins. For any basin it is possible to extract sub-basins down to the level of the drainage areas for individual arc segments (shown in Figure 9b). The western edge of this basin forms part of the drainage divided between the larger Mackenzie (to east) and Yukon (to west) river basins. In this scene along the northern edge of the basin a few artefacts of the triangulation and facet generation can be seen as small divots along the basin boundary. Generally, when the network is dense such artefacts are less prominent and a smoother basin boundary is obtained.



(a) Drainage basin with sub-basin

(b) Drainage areas assigned to individual stream segments.

Figure 9: Basin from the Mackenzie network (a) shaded light grey, along with one of its sub-basins, shown in a slightly darker shaded of gray. Drainage areas for individual arcs or stream segments (b).

Finally, Figure 10 shows the extracted coverage for the Mississippi River drainage basin from the NORTH_AM network for North America. Within this basin the drainage areas to individual stream (river) segments are outlined, while white lines represent the actual drainage network.

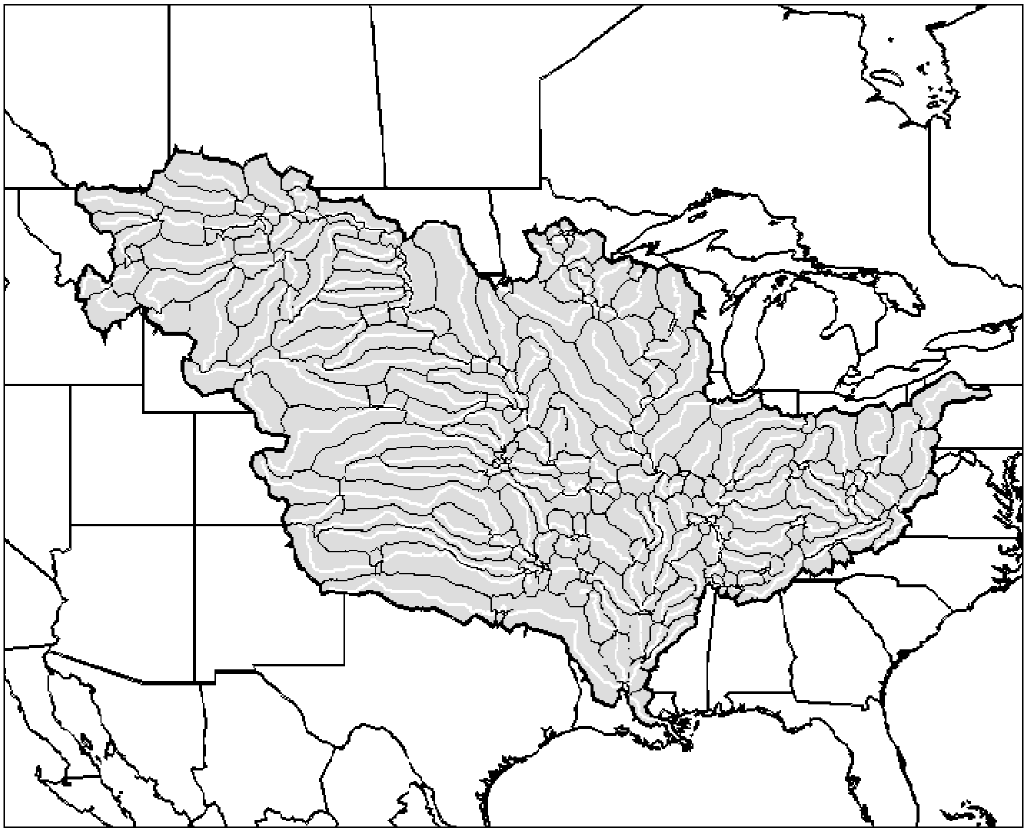


Figure 10: NORTH_AM network, Mississippi River Basin.

5. Discussion and Conclusions

The procedure presented here provides an effective means of drainage basin delineation where drainage network vectors are available and in the absence of an adequately precise DEM. Also for approximation purposes this technique can be favourably compared to DEM based techniques due to its lower computational requirements. Raster based DEMs will often require a significant amount of pre-processing in order to ensure removal of sinks that would interfere with basin delineation procedures. Comparatively editing a vector stream network to ensure proper directionality is quite straightforward and less time consuming. Finally, for large areas with dense drainage networks, which would require a fine spatial resolution, the size of the raster DEMs required to delineate drainage basins would be prohibitively large.

The primary drawback of this technique is the somewhat rough basin boundaries produced when the individual polygons representing drainage to a single arc segment are viewed. However, as the degree of complexity (i.e. larger number of stream segments) and size of the network being extracted increases this becomes relatively insignificant.

The technique can handle very large detailed drainage basins and produce drainage area polygons quickly. The facet generation and assignment algorithm is efficient, and relatively easy to implement. While no attempts were made to extend the technique developed in this study to other problem domains, the basin approximation could be easily used to calculate area drainage for any single arc segment in a drainage network, which could in turn be used to estimate of stream order. Further, the bisection method produces a visually appealing division between parallel and nearly parallel lines and might be used for problems like the generation of lake or double line river skeletons in map generalization.

6. Acknowledgements

Thanks to Rupert Brooks, Rhian Evans, and Dianne Richardson for their assistance in this research.

7. References

Held M (2001) VRONI: An engineering approach to the reliable and efficient computation of Voronoi diagrams of points and line segments. *Computational Geometry* 18(2):95-123

Jones NL Wright SG and Maidment DR (1990) Watershed delineation with triangle-based terrain models. *Journal of Hydraulic Engineering* 116(3): 1232-1251

Mower JE (1994) Data-parallel procedures for drainage basin analysis. *Computers and Geosciences* 20(9): 1365-1378.

Nelson EJ Jones NL and Miller AW (1994) Algorithm for precise drainage-basin delineation. *Journal of Hydraulic Engineering* 120(3): 298-312

Shewchuk JR (1996) Triangle: A two-dimensional quality mesh generator and Delaunay triangulator, Version 1.3. Available from: *http://www.cs.cmu.edu/~quake/triangle.research.html.*