Jehu Peters and Don Metz

Using Graph Theory to Understand First Nations Connections

Children of the Earth High School in Winnipeg, Manitoba, is a school dedicated to incorporating Aboriginal perspectives into all areas of the Manitoba curriculum and is attended by an almost exclusive Aboriginal population. Many of these students come from a First Nations community and are related to others who come from such communities, the majority of which are found in the northern regions of Manitoba.

During a unit on relations, we used as an example an arrow diagram that connected the students to the schools that they had previously attended. This simple exercise caused many students to wonder what other kinds of relationships could be analyzed. After a brainstorming session, the class wanted to know which Aboriginal communities their fellow students most closely associated with, and the class came up with the idea of surveying the entire school to collect the necessary data.

The evolution of computing in the twenty-first century has led to the collection of large amounts of data for business, marketing, and government purposes. *Data mining* is a catchall term to describe the processes used in the collection and analysis of "big data." Typically, visualization is used to help reveal hidden patterns in the data not previously obvious. Our form of data mining led us to understand the basic concepts of pattern discovery and clustering and raised many questions for further study.

The students began by framing the question, "What First Nations communities do you come from?" Many students have parents or grandparents from a number of different communities (some as a result of blended families), so, for purposes of this survey, individual students could consider themselves to "come from" multiple communities. In this way, we were able to illustrate the connectedness of the students in the school even though most of them came from northern, often remote, communities.

The students surveyed 110 students and found that 47 of the 63 First Nations communities in Manitoba as well as six First Nations communities from other provinces were represented at Children of the Earth High School. The students learned a number of useful survey skills, including how to create clear questions, enter data into a spreadsheet, and present their results. The most interesting insights came when we used the framework of graph theory to analyze the data.

GRAPHING WITH GEPHI

Quite simply, a graph is a set of vertices joined by lines called edges. A graph can be used to visually represent a variety of relationships, highlighting data organization, patterns, networks, and connections. We decided to visualize the graph with the help of Gephi, an open-source graph visualization software available online at http://gephi.github.io/. To begin, download and install the Gephi program. The Gephi program reads data in a comma-separated values format (.csv file format) as source data and target data. Students entered the data into a spreadsheet with their name (**id**) in the first column and their community in the second column (see **fig. 1**). Be sure to label column A as Source and column B as Target and save the data as a .csv file (use **File Save As** and select **.csv** in the file format selection box). To capture the fact that one student might have a relationship with more than one community, his or her name could be listed more than once in the first column and matched with a different community in the second column.

To make your graph, open the Gephi program and start a new project. Use **File Save As** to save your project in a folder on your computer. By default, the

Edited by Andrew Benzing

Technology Tips offers a forum for innovative uses of technology in the teaching and learning of mathematics; it appears in every issue of *Mathematics Teacher*. Manuscripts for the department should be submitted via http://mt.msubmit.net. For more background information on the department and guidelines for submission, visit http://www.nctm.org/mtcalls.

Department editors

Andrew Benzing, *abenzing@wssd.org, Strath Haven High School, Wallingford, PA;* ***Heather Lynn Johnson,*** *heather.johnson@ucdenver.edu, University of Colorado Denver; and* ***Steve Phelps,*** *sphelps@madeiracityschools.org, Madeira High School, Cincinnati, OH*

	A	B
1	Source	Target
2	S1	Norway House
3	S1	Poplar River
4	S103	Sayisi Dene
5	S104	Fairford
6	S104	Lake St. Martin
7	S104	Pinaymootang
8	S105	Skownan
9	S109	Brokenhead

Fig. 1 Spreadsheet data link students (Source) with communities (Target).

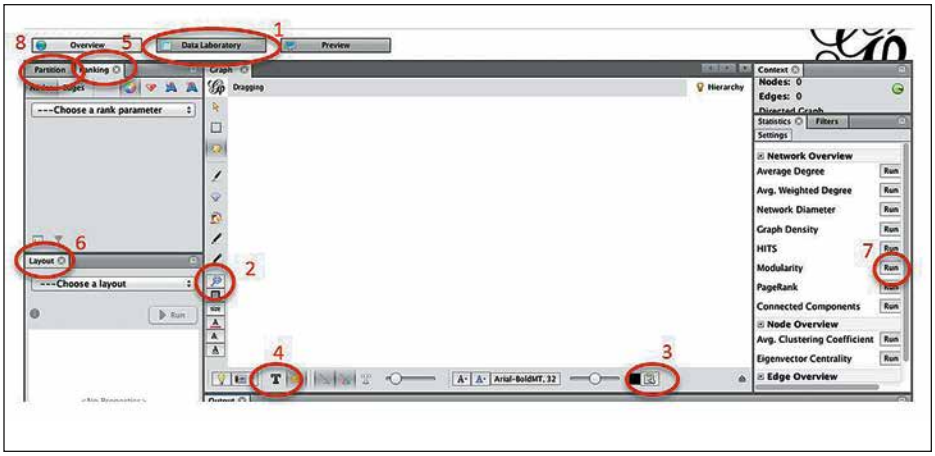

Fig. 2 A screen overview shows options using Gephi.

Overview screen is active (see **fig. 2**). Next, select **Data Laboratory** (labeled 1 in **fig. 2**) to open a new screen and then select **Import Spreadsheet** in the new screen. You can browse to find your .csv file; then select the **As Table** button and choose the **Edges Table** option. Select **Next**; the import defaults are fine, so select **Finish**. Your data should now appear in the **Data Laboratory** window.

Return to the **Overview** screen, and your graph should appear. Gephi selects a layout randomly, but you have access to a wide range of options to modify the layout to suit your own needs. Use the mouse wheel (or two-finger swipe on a track pad) to zoom in and out, and use the right mouse hold **Command** to pan. There is also a centering icon (labeled 2 in **fig. 2**) that is handy.

A wide variety of options to modify the layout of your graph is available. Here are a few simple ones to get you started. It is helpful to see the labels of your nodes. To do this, select the **Attribute** tool (labeled 3 in **fig. 2**), select **Id**, unselect **Label**, and press the **Show Node Labels** button (labeled 4). It is useful to have the node size reflect the popularity of the nodes. To adjust the node size, choose **Ranking** (labeled 5), the **Gem** icon, and choose a rank parameter. Select **Degree** and try 1 for minimum and 50 for maximum.

The **Layout** option (labeled 6 in **fig. 2**) is used to obtain a graph that is easy to read and display. Select **Choose a Layout** for the options. We used the **Force Atlas** ranking parameter with 6500 for repulsion strength and 300 for gravity. Be sure to select **Run** to initiate and **Stop** to terminate. In the same option list, you can run the **Label Adjust** layout to ensure that no labels are touching. You can also use the hand tool to grab any node and drag it where you want.

The modularity algorithm (labeled 7 in **fig. 2**) is used to detect First Nations communities within the graph. Run **Modularity** with default parameters. The graph will describe the number of communities and the size of each. To color the nodes according to the students' community, choose **Partition** (labeled 8), click on the green refresh arrow, select **Choose a Partition**, and then choose **Modularity Class**. Press the **Apply** button.

Finally, select **Preview** (the button next to **Data Laboratory**), and in the **Presets** choose **Default Straight**; you may have to refresh the screen. You can export your graph by selecting the **Export** option in the bottom-left corner of the screen.

VISUALIZING CONNECTIONS

The data were translated into a graph with these features:

- The vertex set contains all the students and the First Nations communities.
- The edge set contains the connections between students and communities.

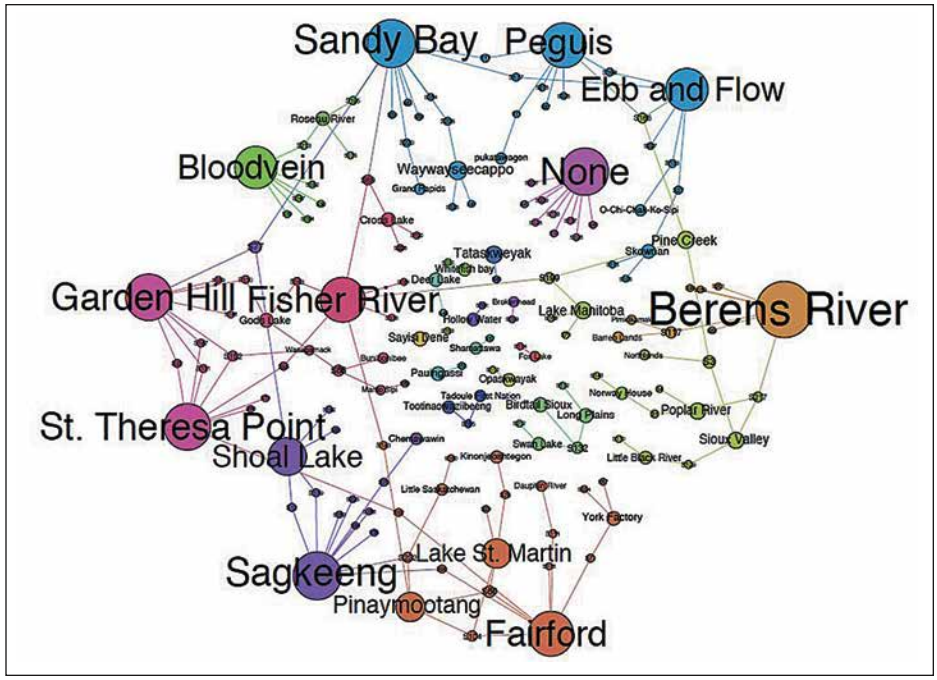

Fig. 3 The graph shows First Nations communities linked by individual students.

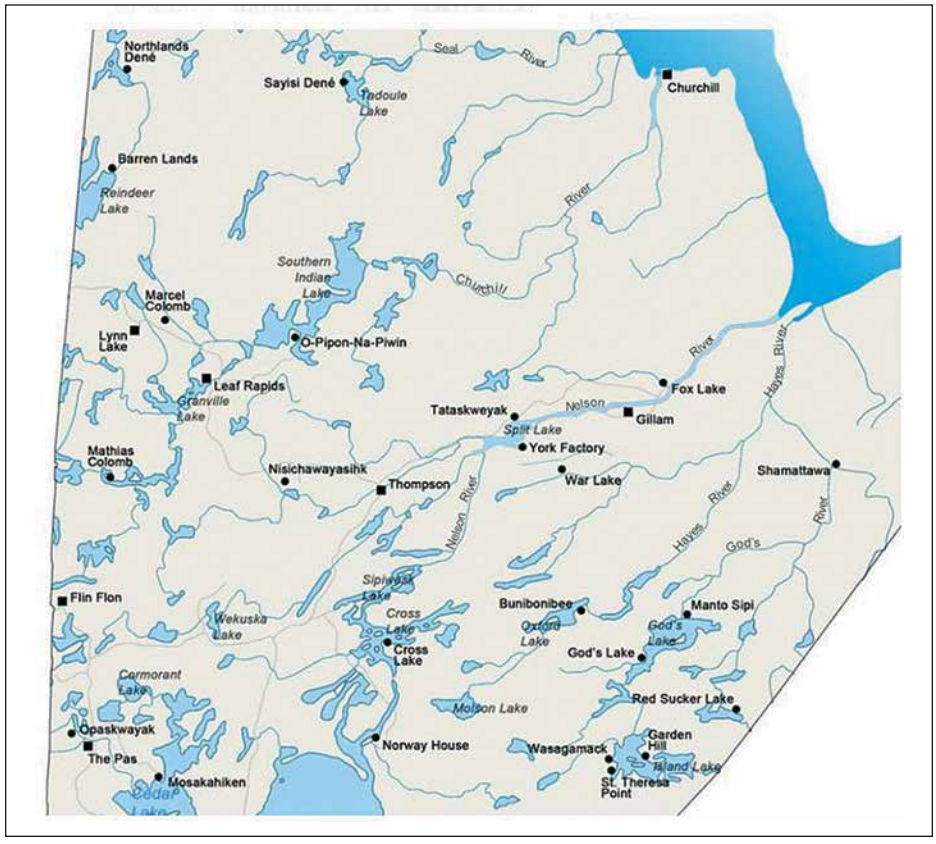

Fig. 4 The geography of northern Manitoba provides insight into community connections.

- Two vertices are connected if a student is connected to the community.
- Because students relate to their communities and their communities relate to them, the graph is undirected.

Our Gephi visualization is shown in **figure 3**. The vertex size indicates how many students come from each First Nations community (the vertex degree).

The color indicates which communities are closely linked (the modularity). By examining the graph, we see that the graph consists of one major component, which contains 76 percent of all vertices, along with a few relatively small components. This major component is somewhat unusual given the low density (6 percent of all possible connections) of the graph. We see that only eleven students with some First Nations community connection are not connected to this main component, illustrating the high degree of connectivity that Aboriginal people often exhibit.

In graph theory, graphs may have subgraphs that can reveal local patterns. For example, if we consider the subgraph for the communities of Garden Hill and St. Theresa, we find direct connectivity for several students even though the communities are located on opposite sides of a lake. However, when we refer to geographical maps, we find that the communities have links by water access and winter roads to one another and to a few other smaller communities, such as Wasagamack and God's Lake (see **fig. 4**). In addition, we find that most of the unconnected reserves come from very northern reserves, such as Shamattawa, Fox Lake, and Sayisi Dene.

In a classroom setting, these insights could be discovered by the students. For example, the following questions could help guide the students in uncovering the geographical patterns:

- In considering the graph, which First Nations communities appear to have direct connectivity?
- How is it possible that the graph indicates that students living on opposite sides of the lake have direct connectivity?
- How would these communities establish such connections in the winter across a lake?
- Why are some communities not connected?

POSSIBLE ADAPTATIONS

This type of project is engaging, encourages student ownership, allows for the introduction of advanced mathematical concepts, and can be used in a variety of situations. Students could formulate research questions, select and use appropriate methods to analyze data, and develop inferences and predictions based on the data. To get started, here are some sample questions that students could investigate:

- Is there a pattern in the previous schools that the students attended?
- What does the graph of a fictional character's relationships look like?
- How do words in a specified text relate to one other according to proximity?
- Do certain student interests (e.g., cars and rock music) cluster together?

We believe that the students responded much better when their examples were personal. Our students appreciated this project on their heritage. As one student commented, this graph is "a good way to show our youth how connected we all are."

JEHU PETERS, jehupeters@gmail.com, is a fifth-year teacher candidate in the Faculty of Education at the University of Winnipeg. He implemented this lesson in his teaching practicum. **DON METZ,** d.metz@uwinnipeg.ca, is a member of the Faculty of Education at the University of Winnipeg. He is Peters's professor and supervisor in physics and mathematics education, and they are working in collaboration with the Inner City Science Center at Niji Mahkwa School in Winnipeg.