

NUMB3RS Activity: The Graph Tells the Story **Episode: "Backscatter"**

Topic: Interpreting Graphs

Grade Level: 8 - 12

Objective: Interpret graphs

Time: 15 - 20 minutes

Introduction

In "Backscatter," Don and his team bust a group of Internet hackers associated with the Russian mob. The mob retaliates, and Charlie must help solve the Internet scheme before the situation worsens. Charlie uses a technique called backscatter analysis to help track the flow of distribution of an Internet attack. Charlie uses the analysis to figure out the prevalence of denial-of-service attacks in the Internet. He gathers data to assess the number, duration, and focus of attacks, and to characterize their behavior.

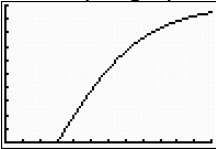
The analysis makes use of the fact that the Internet is a "packet-exchange" system and works much like the distribution of letters through a post office sorting machine. Charlie assumes that when an Internet attack happens, attackers commonly spoof the source IP address field to conceal the location of the attacking host (the source of the attack is hidden by relaying messages through other routers). However, because these messages are happening in real time, a computer can effectively "sample" the activity through a complicated process and ultimately track down where the attack may have originated. The focus of this activity is to interpret some graphs that Charlie might have used in his analysis.

Discuss with Students

This activity will present graphical representations and ask students to interpret the meanings of the graphs. Suppose Charlie wants to graph the cumulative number of Internet messages a person receives in one 24-hour period.

1. Identify the independent and dependent variables.
2. What would be reasonable scales for the x-axis and the y-axis?
3. Sketch a possible graph that would describe this situation.
4. Explain how you could use the graph to predict the cumulative number of messages received after 12 hours.
5. Explain how you could use the graph to predict how long it takes to accumulate 30 messages.

Discuss with Students Answers:

1. The independent variable is the number of hours. The dependent variable is the cumulative number of messages received between time 0 and time x .
2. Answers will vary. A reasonable scale for the x -axis is to go from 0 to 24 hours with a scale of 2 hours, and a scale for the y -axis would be from 0 to 50 with a scale of 5
3. Sample graph:

4. Locate $x = 12$ on the x -axis. Go up until you intersect the graph and read the corresponding y -value.
5. Locate $y = 30$ on the y -axis. Go over until you intersect the graph and read the corresponding x -value.

Student Page Answers:

1. The time durations on the x -axis range from 1 minute to 7 days. Because this is such a large range of values, he uses three different units on one axis: minutes, hours, and days. Therefore, the scale marks on the graph represent different units.
2. The percentages on the y -axis range from 0 to 100, but a logarithmic scale is used: the distances from 0 to 1, 1 to 10, and 10 to 100 are all the same. Each y represents the percentage of attacks with time duration less than or equal to x .
3. approximately 8%, 20% and 100%
4. approximately between 1 and 3 minutes
5. about 3%
6. about 1.5 minutes, 2 minutes, and 30 minutes
7. No, because the chance of an attack lasting 7 days is very close to 0.
8. about 3 minutes to 30 minutes

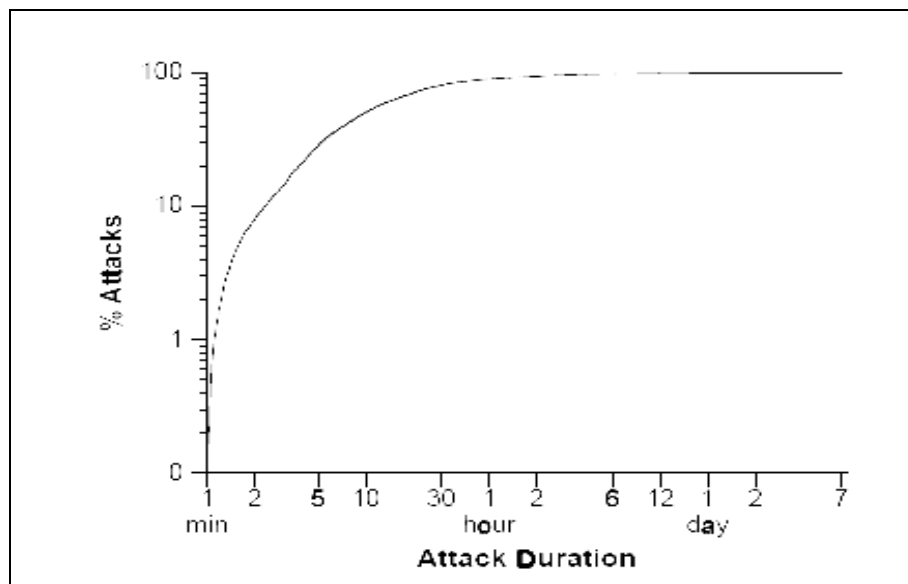
Name: _____

Date: _____

NUMB3RS Activity: The Graph Tells the Story

In "Backscatter" Don and his team bust a group of Internet hackers associated with the Russian mob. The mob retaliates, and Charlie must help solve the Internet scheme before the situation worsens. Charlie uses a technique called backscatter analysis to help track the flow of distribution of an Internet attack. Charlie uses the analysis to figure out the prevalence of denial-of-service attacks in the Internet. He gathers data to assess the number, duration, and focus of attacks, and to characterize their behavior. The process is very complicated.

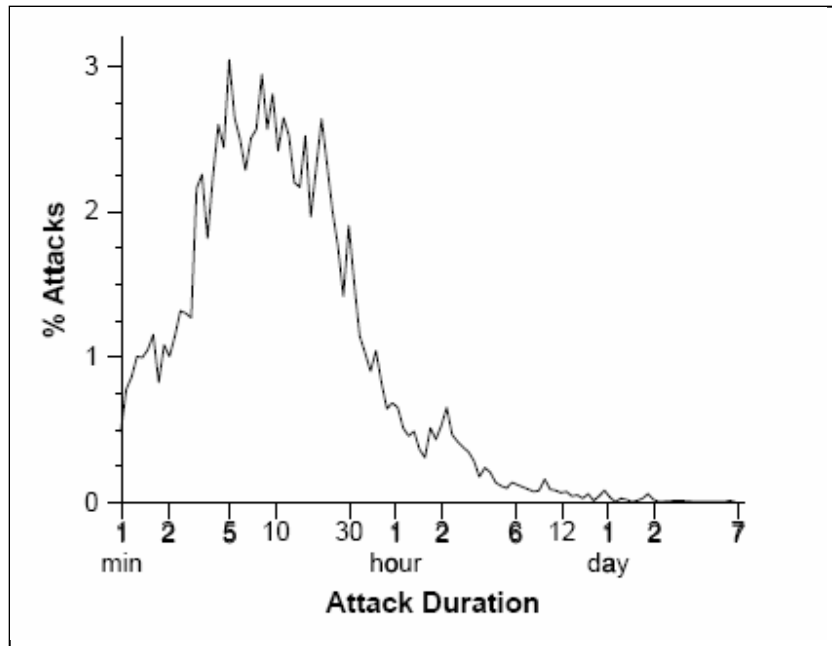
Suppose that Charlie has finally produced two graphs shown below and needs to interpret them. The list of attacks on Don's Internet provider over the course of one week is first sorted into increasing order according to the length of time or duration of each attack. The first graph shows the cumulative distribution of these attack durations. The second graph shows the probability density of attack durations for Don's Internet provider over the course of one week.



[Source: <http://www.caida.org/publications/papers/2001/BackScatter/usenixsecurity01.pdf>]

1. Describe the units on the x-axis. Explain how Charlie can use three different units to label the same scale. (This scale is called a logarithmic scale. For more information about logarithmic scales, see the Extensions page.)
2. Explain the units on the y-axis. Why do you think the units are written in this way? (This scale is also a logarithmic scale.)
3. According to the graph, what is the cumulative % Attacks for an attack duration of 2 minutes or less? 10 minutes or less? 12 hours or less?
4. According to the graph, describe the time durations of the shortest 10% of the attacks.

This graph shows the probability density of attack durations for Don's Internet provider over the course of one week. The y-axis is the percentage of attacks that lasted a given amount of time.



[Source: <http://www.caida.org/publications/papers/2001/BackScatter/usenixsecurity01.pdf>]

5. What percent of attacks lasted 5 minutes?
6. If the % Attack was 1%, what are the possible Attack Durations?
7. If a mob member claimed that an attack lasted for 7 days, would you believe it? Why or why not?
8. During which time interval do all the attack durations have % Attack greater than 2%?

The goal of this activity is to give your students a short and simple snapshot into a very extensive math topic. TI and NCTM encourage you and your students to learn more about this topic using the extensions provided below and through your own independent research.

Extensions

Introduction

Reading and interpreting graphs is an important skill, as graphs are everywhere in our information driven world. It is important to understand the scale and the meaning of the axes and to determine whether the graph is a graph of values, rates, or accumulations. Graphs are often used to analyze the nature of changes in relationships.

For the Student

- Find a graph in a newspaper or other resource and describe all of the information contained in the graph.
- The first graph in the activity is graphed using a log-log scale. Redraw the graph using a standard scale. Describe why the log-log scale is preferable. To learn more about the logarithmic scale and its applications visit this Web site: http://en.wikipedia.org/wiki/Logarithmic_scale.
- Learn more about backscatter analysis and how it is used to trace Internet attacks. To read the article that was used for the episode "Backscatter" and to view the graphs used in this activity, visit the Web site below. <http://www.caida.org/publications/papers/2001/BackScatter/userixsecurity01.pdf>
- An educational video describing backscatter analysis can be downloaded from the Web site below. <http://www.caida.org/publications/animations>

Additional Resources

Visit the Web site below for more resources on reading and interpreting graphs. The lesson is designed to introduce students to graphing functions and to reading simple functions from graphs. Many of the examples are motivated by situations described by the graphs.

<http://www.shodor.org/interactivate/lessons/fm4.html>

The National Curve Bank is a resource for students of mathematics that provides geometric, algebraic, and historical aspects of curves through animation and interaction. Topics include calculus, pre-calculus, history of mathematics, limits, and fractals.

<http://curvebank.calstatela.edu/home/home.htm>

Visit this Web site for five activities that involve reading and interpreting graphs.

<http://www.nottingham.ac.uk/education/shell/graphs.htm>