

THIS MAN COULD RULE THE WORLD

How Albert-László Barabási went from mapping systems to controlling them

BY GREGORY MONE

In 1736 the Swiss mathematician Leonhard Euler ended a debate among the citizens of Königsberg, Prussia, by drawing a graph. The Pregel River divided the city, now Kaliningrad, Russia, into four sections. Seven bridges connected them. Could a person cross all seven without walking over the same one twice?

Euler began with a map that cleared away everything—the homes and streets and coffeehouses—irrelevant to the question at hand. Then he translated that map into something even more abstract, a depiction not of a physical place but of an interconnected system. The four sections became

dots, and the seven bridges became lines. By transforming Königsberg into simple nodes and edges (as mathematicians have come to call such abstractions), Euler could subject the system to mathematical analysis. In doing so, he proved that a person could not cross all seven bridges without walking over the same one twice. More important, he mapped a network for the first time.

Over the next two centuries, scientists built on Euler's work to develop graph theory, a branch of mathematics that would eventually serve as the basis for network science. But it wasn't until 1959—when the Hungarian mathematicians Paul Erdős and

Alfréd Rényi proposed a means by which complex networks evolve—that a defined theory of networks began to emerge. And it was only in the mid-1990s that scientists began to apply that theory to really complex problems. Before then, large data sets were difficult to obtain and even more difficult to process. But as data became more accessible and processing power cheaper, researchers began applying graph theory to everything from protein interactions to the workings of the power grid.

Albert-László Barabási, a Romanian-born physicist at the University of Notre Dame, was one of those researchers. In the past decade and a half, he has transformed the way his colleagues understand networks at least twice. His theories have influenced important developments in engineering, marketing, medicine and spycraft. And his research may soon allow engineers, marketers, doctors and spies to not just understand and predict network behavior, but also to control it.

In the beginning, though, Barabási, like Euler, was mostly interested in mapping complex systems. He was particularly interested in the Erdős-Rényi model, which held that complex networks were random, and if they grew large enough each node would have roughly the same number of links as any other node over time. In 1998, Barabási and his students at Notre Dame saw an opportunity to study the implications of that theory on a really big data set: 325,000 pages from Notre Dame's Web domain. When they ran the numbers, nearly all the pages did in fact have about the same number of links. But a few dozen were different. They had upward of 1,000 incoming links. At the time, Google's PageRank was already exploiting this quality to produce more-relevant search results, but to network theorists the notion was radical and had implications far beyond the Web. Barabási later wrote that "we caught a glimpse of a new and unsuspected order within networks, one that displayed an uncommon beauty and coherence."

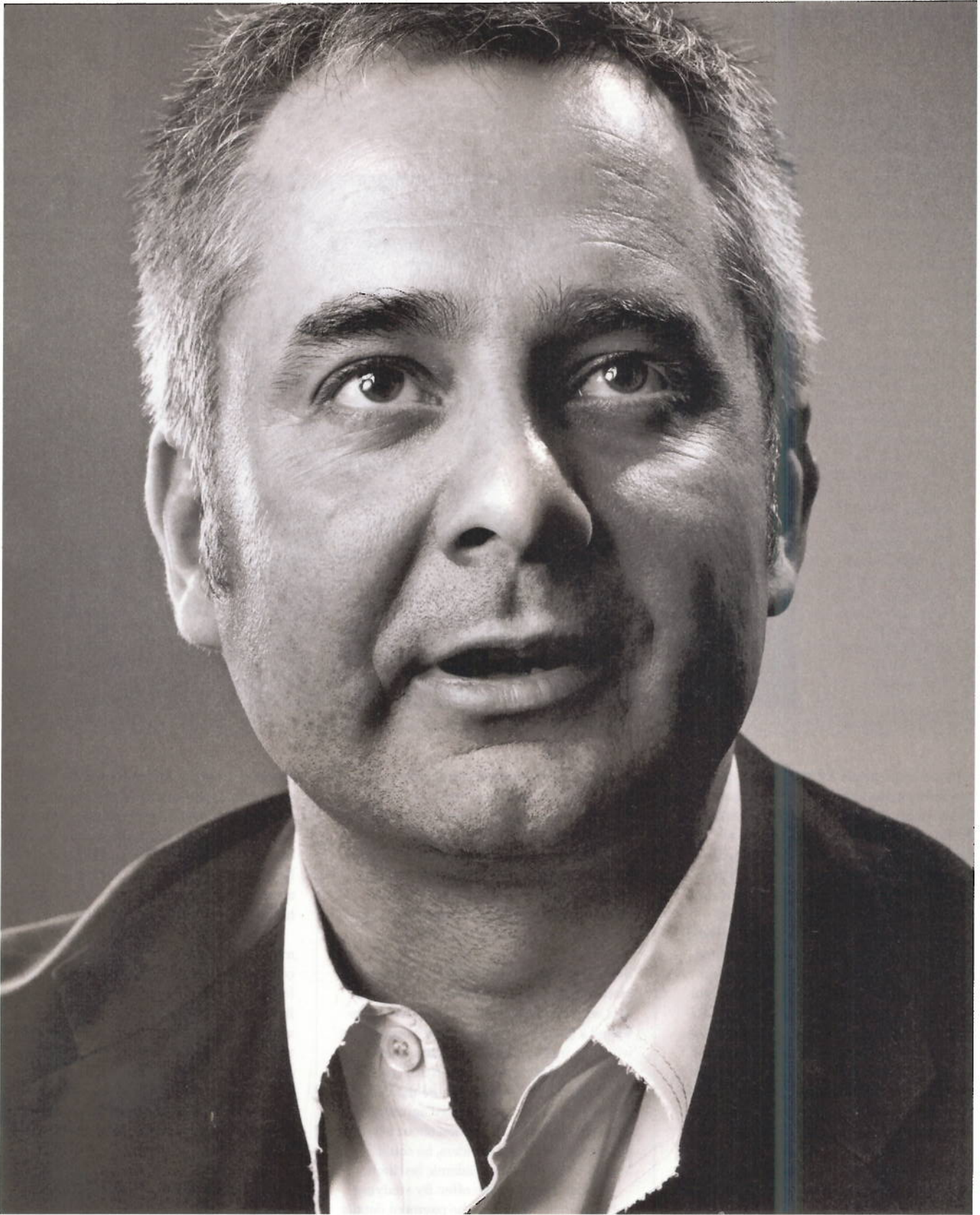
Faced with a contradiction between the Erdős-Rényi model and his findings, Barabási mapped several other large and complex systems, including the connections between transistors on computer chips and the collaborations between actors in Hollywood. In each case, highly linked nodes, which he called hubs, were the defining characteristic of the network, not just an anomaly but an

→ 1877 Thomas Edison invents the phonograph.



→ 1878 By an act of Congress, the collection of morbidity data on cholera, smallpox and yellow fever begins.

JULIAN DUFORT; FACING PAGE: HULTON ARCHIVE/GETTY IMAGES

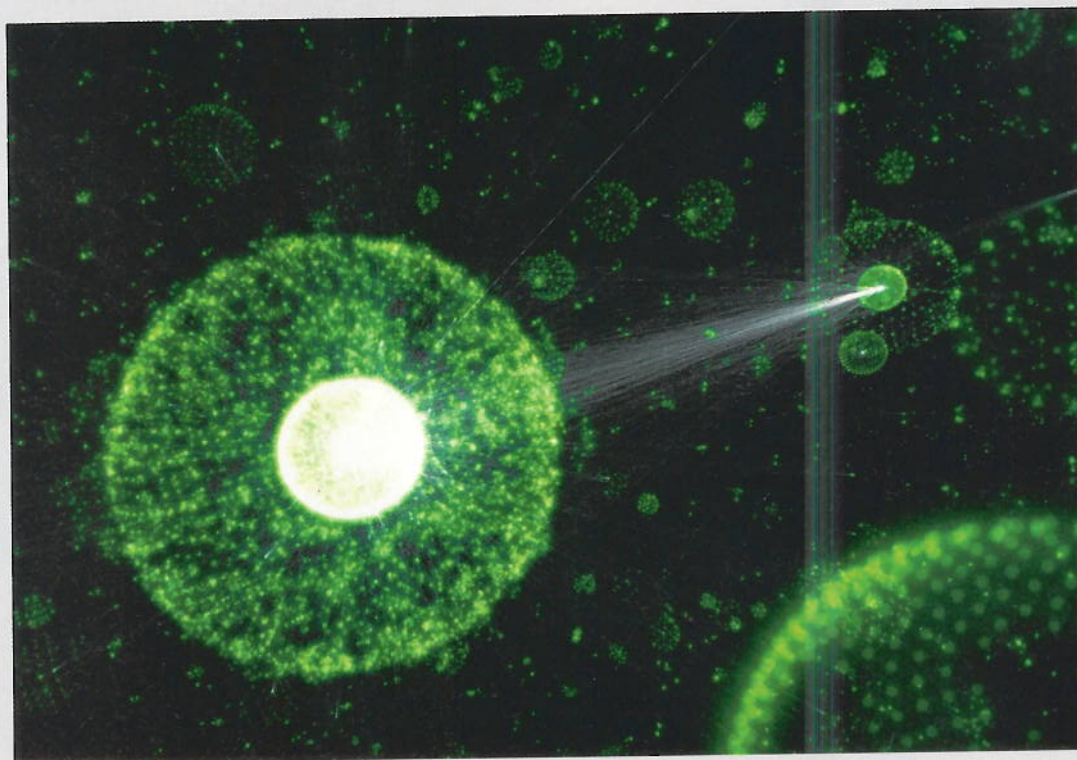


→ **1895** Belgian lawyers Paul Otlet and Henri La Fontaine begin to collect index cards of information to answer factual questions by mail. By 1934, they had amassed 15.6 million cards.

→ **1935** George Gallup founds the American Institute of Public Opinion and begins the first rigorous collection of opinion polls.

A UNIVERSE OF HUBS

To illustrate how hubs act as an organizing principle within complex networks, Mauro Martino, a computer scientist and interactive designer in Albert-László Barabási's lab, plotted 325,729 Web pages in the University of Notre Dame Web domain [green nodes]. He also mapped the 1,497,134 links that connect those pages [white edges]; for clarity, he showed only the strongest connections. Nodes with many connections are hubs. Less-connected nodes cluster around them like planets gather around a star.



organizing principle for engineered and natural systems alike. With his student Réka Albert, Barabási updated the Erdős-Rényi model to reflect the existence of hubs in real-world networks. In doing so, he created a tool for scientists to map and explore all manner of complex systems in ways they had never thought to before.

Barabási's paper on hubs quickly evolved into one itself, becoming one of most cited in the field of network science. He turned it into a popular book, *Linked*, and later got his own lab at Northeastern University. Scientists in other fields began to draw on hub theory. Cancer researchers used it to better understand how a network of proteins suppresses tumors in the body. Biologists, aided by Barabási, used it to determine antibiotic targets within the metabolic networks of drug-resistant bacteria; the research could provide an entirely new avenue for drug discovery. There are even signs, Barabási says, that the intelligence community is

using his work to map terrorist networks. "It's a matter of wording," he says. "There are lots of little hints that they are using it."

But the translation of his insights into applications did not hold Barabási's interest for long. He is a theorist, not an applied scientist. And once he had the ability to map a system, he says, his next challenge was to predict its behavior.

BARABÁSI GOT HIS CHANCE to work on prediction in 2006. That year, a man called him with an unusual offer. He said he represented a European mobile-phone consortium, which he insisted remain unnamed, and he possessed an intriguing trove of data: the anonymized records of more than six million subscribers. If Barabási agreed to mine the data for information about why customers switched providers, he could also use it for his own academic research.

Barabási accepted the offer. By studying patterns in call logs and the payment details

attached to each number, he and the members of his lab were indeed able to construct an algorithm that identified customers who were likely to switch providers. In exploring the data, though, he also found that it identified the cellphone towers that subscribers accessed when making calls, which allowed him to gauge the physical location of callers.

Physicists have been predicting the movement of particles and planets for centuries, but they had never successfully forecast the comings and goings of people. Barabási and physicist Chaoming Song, also at Northeastern, hypothesized that if they treated those callers as particles, they could predict a person's location at any given time. They wrote software to map the movements of 50,000 callers. Each cell tower became a node. When a user traveled from one node to another, the path was marked by an edge. They then derived each individual's entropy, which measures the degree of randomness or uncertainty in a

MAURO MARTINO/BARABÁSI LAB

→ **1936** English mathematician Alan Turing proposes the stored computer program.

→ **1936** The U.S. federal government issues the first Social Security numbers.

→ **1940s** Electronic computers allow computations of all kinds to be automated with increasing efficiency.

system. By combining movement data with entropy figures, Barabási and Song found that they could predict a person's location, within a square mile, with up to 93 percent accuracy. No one, not even those who traveled frequently outside their usual circuit, was less than 80 percent predictable.

Researchers are only just beginning to integrate Barabási and Song's findings into real-world applications. Epidemiologists use airline travel data to predict the spread of disease from city to city. Barabási and Song's findings could allow them to home in on a single block. Predicting how, when and where people move could help traffic engineers find ways of easing congestion and urban planners design cities in the developing world capable of handling large inflows of migrant workers. In 2009, Barabási and several students used prediction algorithms to explain why mobile-phone viruses aren't prevalent now but could be a serious threat as soon as enough phones are governed by a single operating system.

Predictive science does have a downside. After publishing his work, Barabási received a flood of e-mails accusing him of opening the door to Big Brother applications. Authorities could use his algorithms paired with the GPS data collected on mobile phones to track and predict the movement of citizens with remarkable precision. And if people could predict behavior within a system, might they not also find a way to control it?

IN 2009, WHEN HE was visiting the University of Minnesota to present his work, control was much on Barabási's mind. Shortly before the lecture, he got into a conversation with an engineer. "After five minutes it became absolutely clear that he had no idea what I did," Barabási says. "So I asked, 'What do you do?' And he said, 'I'm a control theorist.'"

Engineers use control theory to predict how systems will respond to various inputs, which in turn helps them make robots that can catch baseballs, cars that take sharp corners with ease, and planes that don't fall from the sky. Barabási had never heard of it, so his new friend led him to a whiteboard and drew out the basic equations. Barabási noticed how similar they were to the ones he used to map networks, and he decided to incorporate them into his own work.

Like prediction, control required evaluating an object as a system with nodes of

varying importance. A car for instance: "It is made of 5,000 components," Barabási says, "yet you as a driver have access to only three to five nodes"—the steering wheel, the gas pedal, the brake, and maybe the clutch and shifter. "With those three to five knobs, you can make this system go anywhere a car can go." What he wanted to know was if he could look at *any* network, not just engineered ones, and find those control nodes. Among the thousands of proteins operating within a cell, could he find the steering wheel, the gas pedal and the brake?

Barabási asked Yang-Yu Liu, a physicist

be useful, like a car with 3,500 steering wheels, but Barabási points out two things: Whereas the neuronal map of *C. elegans* is complete, scientists have determined only about 5 percent of the connections in the yeast cell's gene network. The more data scientists feed into the model, the better they can map connections in the network and the fewer control nodes they might need to operate the system. "We know these maps are incomplete," Barabási says. "But they're getting richer every day." He also says his theory applies to total control of a network. Scientists

AMONG THE THOUSANDS OF PROTEINS OPERATING IN A CELL, COULD BARABÁSI FIND THE STEERING WHEEL, THE GAS PEDAL AND THE BRAKE?

in his lab, and Jean-Jacques Slotine, a control theorist at the Massachusetts Institute of Technology, to help him locate "control nodes" within networks. Control nodes take instructions or signals from outside the network (for example, a foot on the gas pedal) and transmit them to nodes within the network (the fuel-injection system). To find them, Liu borrowed an algorithm, developed by Erdős and Rényi five decades prior, that acts as a signal moving through the network. It starts at one node and follows a random edge to another node, at which point it "erases" every other edge but the one it came in on and the one it will go out on. The algorithm runs through the entire network over and over until it finds the minimum set of starting points needed to reach every node in the system. Control these starting points, and you control the entire network.

The group tested the algorithm on 37 different networks, including a constellation of alliances within a prison population, the metabolic pathways in yeast, and several Internet communities, including Slashdot and Epinions. They found that denser, more interconnected networks tended to have fewer control nodes per capita. For instance, the brain of the highly studied worm *C. elegans*, a network of 297 neurons, has only 49 control nodes. The network of genes operating in a yeast cell produces 4,441 proteins, but Barabási found that he would need to control 80 percent of them, or 3,500, to control the system.

This sounds like too many points to

who want partial control—say, to elicit a particular protein expression within a cell—would need to master far fewer nodes.

As with most of Barabási's work, it will take time to make it useful. Finding the points of control is one thing. Actually exerting influence over a given network, be it Facebook or the human immune system, is an entirely different challenge.

The first breakthroughs will most likely take place in medicine. By identifying control nodes in cell growth systems, scientists could return mature cells to their embryonic state, creating a new source of stem cells. "Some diseases are all about lack of control," Barabási says. "If you were able to gain control over them at the cellular or neuronal level, you might be able to cure the disease."

Control can be used for ill as well as good, of course. Marketers could learn how to better manipulate consumers, and governments could develop new techniques to cow citizens. It's up to us, Barabási says, to define how control should be applied and how it shouldn't be. "What we have to realize is that control is a natural progression of understanding processes," he says. "But control is a question of will, and will can be controlled by laws. We have to come together as a society to figure out how far we can push it."

Gregory Mone is a contributing editor at POPULAR SCIENCE and the author, most recently, of the novel Fish.

→ **1945** American engineer Vannevar Bush speculates on the idea of a "Memex" device that would provide computerized access to the world's knowledge—a precursor to hyperlinks.



→ **1953** James Watson and Francis Crick determine the physical structure of DNA.