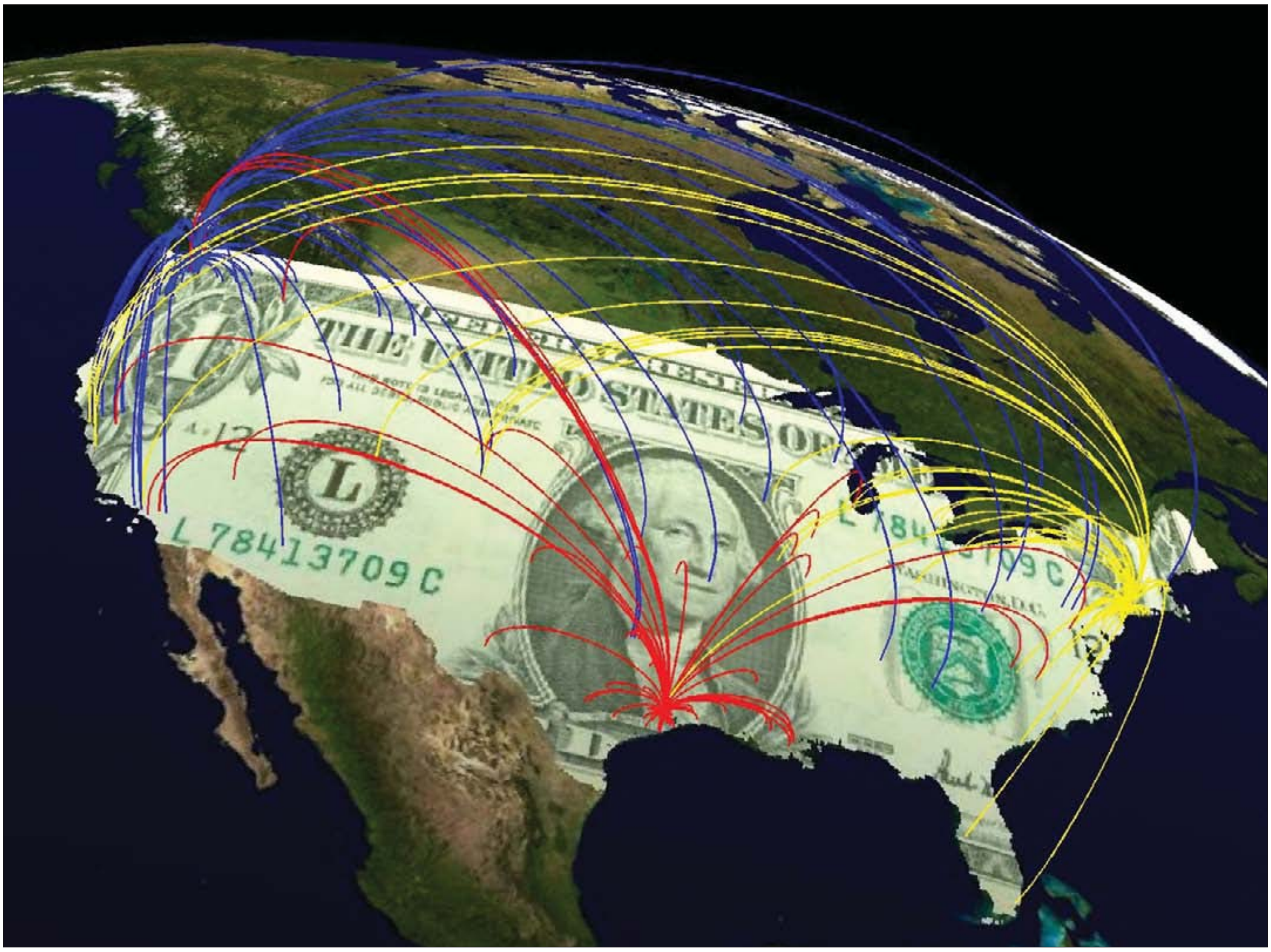


Following the money

Tracking dollar bills and hunting for hidden treasures may seem like unusual pastimes for a physicist. Yet as **Dirk Brockmann** describes, data derived from such pursuits can reveal fascinating patterns in how people travel – and even hint that human mobility could be governed by fundamental physical laws

In 1271 a 17-year-old boy called Marco Polo embarked on an adventure that took him from his hometown of Venice, across Europe, the Middle East and Asia, all the way to Beijing. He returned home 24 years later, having travelled over 24 000 km. A trip of this magnitude was unprecedented in Polo's time; almost 750 years later, he is rightly remembered for it. But what was an exception in medieval Italy is today an everyday occurrence: our 21st-century world is on the move. Every year, more than three billion passengers use the international air-transportation network, flying on a complex web of routes that connects more than 3000 airports worldwide. This global connectivity has made it possible for people to go from virtually any place on the planet to any other within a matter of days.

However, long-distance air travel reflects only a tiny percentage of human mobility. Our world is also covered by a dense network of roads and highways, frequently operating at their maximum capacity. Given the sheer complexity of human-mobility patterns and transportation networks, it may seem bold to suggest that there may be basic underlying principles that govern how such networks evolve, that mobility as a whole may follow fundamental laws, and that some regularities are hidden within the complex way we travel. Yet addressing these questions is of fundamental importance, as increasing human mobility also facilitates the global spread of infectious diseases (see article on page 26). Understanding human mobility is therefore not just interesting, it can actually save lives.

Dirk Brockmann is a theoretical physicist in the Department of Engineering Sciences and Applied Mathematics at Northwestern University in Evanston, Illinois, US, e-mail brockmann@northwestern.edu

Where do you want to go today? The worldwide air-transportation network. The red lines indicate routes with high traffic.

Movement by proxy

But how can one possibly begin to analyse the multitude of traffic and transportation networks in a comprehensive way? The obvious option – monitoring human movement directly – would be costly and invasive, as we cannot simply tag people with radio collars, as biologists do with migrating birds. But the task of compiling mobility data for all possible means of transportation into a massive dataset is not easy either. When my colleague Lars Hufnagel and I started thinking about large-scale analysis of human-transportation networks in 2005, we knew our mission would be very difficult at best, and impossible at worst.

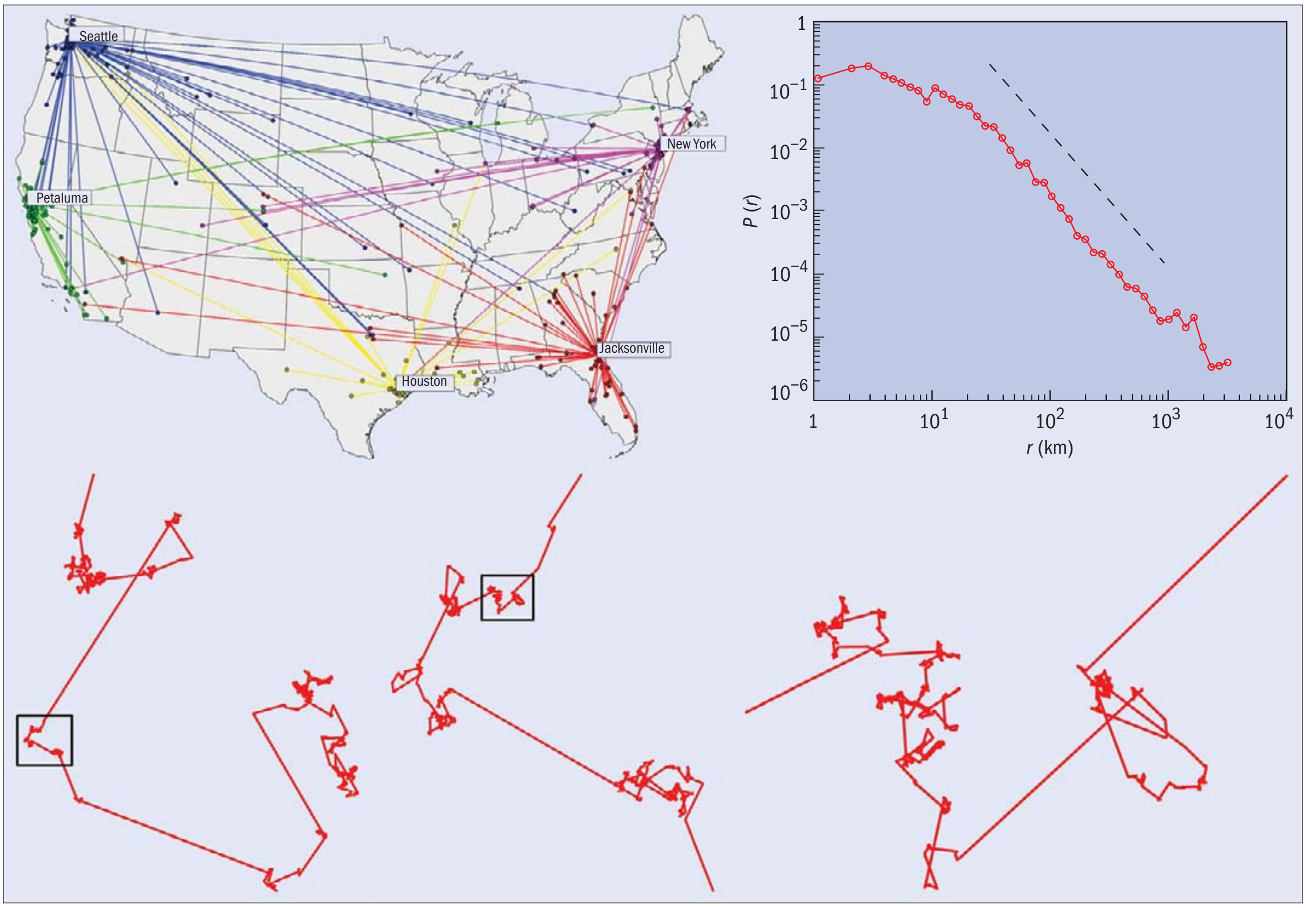
I was still wrapped up in these thoughts when, on my way home from a physics conference in Montreal, Canada, I decided to visit my old friend Dennis Derryberry, who lives in the green mountains of Vermont. One evening during this visit, Dennis asked "So Dirk, what are you working on?". I replied that I was interested in the patterns that underlie human travel, and I described some of my efforts to better understand human mobility. "It's just amazingly difficult to compile all the data," I concluded. Dennis paused a while and then inquired "Do you know the website *Wheresgeorge* [www.wheresgeorge.com]?". I did not, but once he showed it to me it became clear that, in a flash, it could solve a number of our most pressing problems.

Simply put, *Wheresgeorge* is an online game that tracks the geographic circulation of dollar bills, which carry the face of US President George Washington. The way it works is that individual bills are marked by a large community of "Georgers" (the site has over four million registered users) throughout the country. Whenever one of them gets hold of a marked bill, they visit the website and provide their current zip code and the serial number of the bill. Once the bill is back in circulation it can be reported again at another time and place by some other person, thereby generating a trajectory of the bill throughout the country. For each registered bill one can monitor these movements and study the logs posted by individual finders. Forming an image of millions of these dollar-bill journeys in my head, I was convinced that analysing the data could reveal essential properties of human mobility – which is, after all, the driving force behind the dispersal of banknotes.

After I returned from Vermont, Lars and I started extracting the data from the website. We began our analysis by asking a very simple question: what is the probability $P(r)$ that a bill traverses a certain distance r in a short period of time? Common sense and everyday experience suggests that this probability ought to decrease with distance because long journeys are less likely than short ones. But the form of this expected decrease turned out to be surprisingly simple. We found that $P(r)$ falls off with distance according to an inverse power law: $P(r) \sim 1/(r^{1+\mu})$ where the exponent $\mu = 0.6$ (see 2006 *Nature* **439** 462).

This simple mathematical regularity in the geographic circulation was not only a surprise to us, it immediately implied important consequences. One implication is that the successive geographical movements of bank notes are similar to a class of random walks known as Lévy flights (see figure opposite). Unlike ordinary diffusion processes, Lévy flights have a self-similar structure. This means that, just like fractals, their trajectories look similar close up as they do from a distance. Lévy flights are more than just a mathematical peculiarity. They have been discovered in many complex systems ranging from the foraging movements of animal species to the fluctuation of stock prices.

The discovery that travelling banknotes exhibit a Lévy-like pattern has another very important conse-



Show us the money The trajectories of individual banknotes with initial entries in Seattle (blue), Petaluma (green), New York (purple), Houston (yellow) and Jacksonville (red) (top left). The lines connect the initial entry location and locations where bills were reported less than a week later. The probability $P(r)$ of a bill traversing a distance r in a one-week period (top right). The dashed line is a power law as described in the text. The observed power-law relation in banknote dispersal indicates that their spatial trajectories are similar to Lévy flights – random walks that are statistically self-similar, shown here at three successive magnifications (bottom).

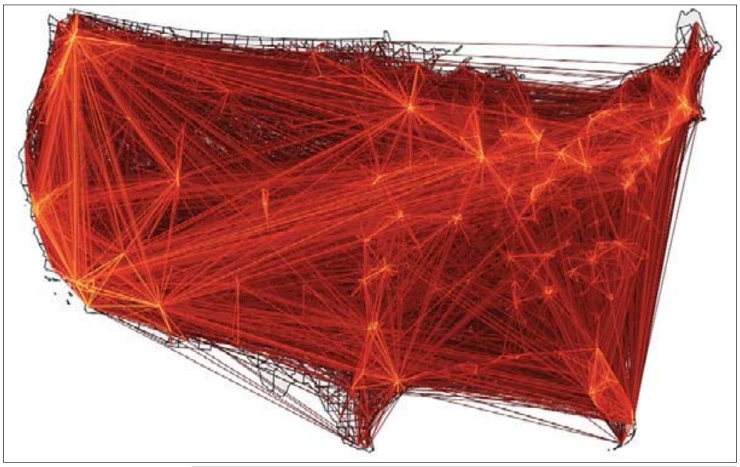
quence: it indicates that modern patterns of human movement are inconsistent with assumptions made for previous models for the spread of diseases. They assumed that infected individuals behaved like diffusing particles. In diffusion, position scales with time in a square-root fashion, $X(t) \sim t^{1/2}$. Lévy flights, in contrast, exhibit the scaling relation $X(t) \sim t^{1/\mu}$, where $\mu < 2$. Such behaviour is termed "superdiffusive" and the equations required to describe these processes are beyond those of ordinary calculus. In fact, the theory of fractional calculus is required to describe them properly.

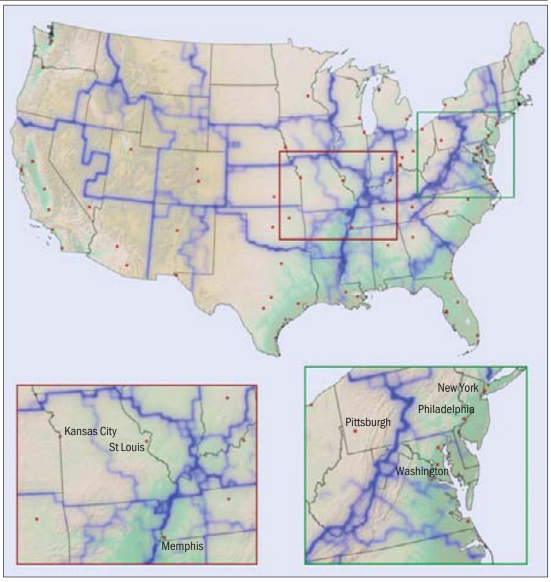
Ties that bind and divide

In addition to scaling laws, the circulation of money can tell a story about the lines that separate us as a society. As communication and travel continue to bring people together on a global scale, it is interesting to investigate the significance of international, national and regional borders. Is it still appropriate to think in terms of the traditional geographical territorial patches that have evolved over decades and centuries? How might these legal borders differ from others that have no standing in law but are more tightly coupled to the way we actually behave – for instance to the way we travel?

It turns out that we can use the flux of several million banknotes in the US to answer these questions. The basic assumption in our analysis is that the more traffic between two places, the more tightly they are coupled and the less "distant" they effectively are. The entire flux of banknotes can be quantified by a network made up of vertices (representing each US county) labelled with an index i and connected by links w_{ij} that measure the number of banknotes travelling between vertices i and j. To establish which geographic areas function as "pseudo-countries", one can group vertices into modules such that the flux is large within the modules and small between them. This is a non-trivial task for large-scale networks, but efficient network algorithms exist that successfully solve it. By applying these to the multi-scale mobility network derived from the flux of money, we recently extracted effective geographic borders that are implicitly encoded in human mobility.

As the figure on page 34 shows, these effective mobility borders coincide with administrative borders only partially, and quite frequently borders exist in unexpected places. The US state of Missouri, for example, is cut in half, with one half belonging to the city of St Louis in the east and the other to the Kansas City area in the west. Further east, the Appalachian mountain range makes up an effective boundary that bears only a loose relation to the various state borders that intersect it. Knowing these effective boundaries may





Unseen boundaries
The flux of banknotes in the US between 3109 counties. This multiscale mobility network encodes a hidden structure of effective geographic borders in the US, shown as blue lines on the map below. Within each blue-bordered region, the flux of banknotes is large, but the number of notes entering or leaving (the external flux) is small. Thicker lines indicate a higher degree of separation between adjacent regions.

in the future allow interdisciplinary scientists to address a number of interesting questions – including to what extent these banknote barriers correlate with social and cultural boundaries.

Mobile phones and treasure hunts

Interesting as they are, banknote studies are by no means the only proxy for gathering data on human-mobility networks. Two years after my colleagues and I first studied the dispersal of dollar bills, Albert-László Barabási and his research team at Northeastern University in Boston published an article analysing the trajectories of individual mobile phones. Barabási's study made use of the fact that mobile-phone companies collect data on where individual phones connect to mobile-phone antennas. As these antennas make up a fine geographic grid, the sequence of antennas that individual phones connect to represents a geographic trajectory of these phones.

The mobile-phone study confirmed the power-law relationship discovered in money circulation, but it was also able to address important questions that the dollar-bill study could not. The fact that people carry their own mobile phones – rather than passing them between each other as they do banknotes – meant that Barabási and his team were able to extract individual mobility networks for each mobile phone in the dataset. This allowed the researchers to examine the underlying reasons why power-law relationships are observed in proxy mobility networks. To do this, they borrowed from physics the well-established concept of entropy, a measure for uncertainty in an unpredictable system, and found that in the mobile-phone mobility network, this quantity was much smaller than expected. This result implied that human mobility is more predictable than we might think.

One additional limitation of the bill-tracking study is that, by definition, it deals only with mobility in the US, and thus is not able to show whether the observed scaling laws hold in other regions of the world. Might not Europeans, travelling over often smaller distances and on a very different transport network, exhibit mobility patterns that differed from those of Americans?

Fortunately, the Internet is a great source of mobility data. In 2006 I read a column in a newspaper about a new type of online game known as "geocaching". This is essentially a treasure hunt based on Global Positioning System (GPS) technology. Players hide treasures, or caches, at various physical locations and publish the GPS co-ordinates on a website. Other players can then download these co-ordinates onto their own GPS device and use this information to locate the cache. If they are successful, they log their find and often exchange toys and gifts at the cache site.

This sounded fun to me, so in an effort to make my daughters Hannah and Lili (then aged eight and five) spend more time outdoors, I bought a GPS device and introduced them to the game. They were excited and so we embarked on our first attempt to locate a cache in the woods near where we live. In this cache we found what is known as a "travel bug", or trackable item. Cachers transport these dog-tagged items from one cache to the next, logging the trace of their mobility patterns online. As geocaching is an international game, and millions of trackable items move between caches, I realized that aspects of human mobility in Europe could be addressed by analysing the trajectories of these trackable items.

After I contacted the people behind the geocaching website, they provided their entire dataset for scientific analysis. Based on this data, I and a colleague Fabian Theis computed the multiscale mobility networks for various European countries as well as the US. Contrary to our expectations, we found that all these networks are remarkably similar in their statistical properties. For instance, the distribution of link strength, w_{ij}, follows the same mathematical law – yet another indication that human mobility is governed by fundamental underlying laws.

Studies of banknote circulation, games such as geocaching, or monitoring the movements of mobile phones are all first steps in a new kind of science. The amount of data on human behaviour that are currently collected online and that can serve as proxy data for what we do is immense. Complexity research, complex network theory and concepts from statistical physics will be one of the most promising scientific combination of tools and methods to deal with these datasets. Together, they form a new kind of science for the understanding of human behaviour at large.