NETWORKS
Bridging our understanding of the world

Innovation and Entrepreneurship
Featuring Tillman Gerngross, Ph.D.
p. 3

Massive Open Online Courses
Education Spread Through Technology
p. 25

Perceptions and Use of File Sharing
Data from Dartmouth Students
p. 40
Dear Reader,

In the 1960s, Stanley Milgram conducted the “small-world” experiment, which measured the average network length between two randomly selected people in the United States. People living in cities in the Midwest were told to send a letter to specific addresses in Boston using only known contacts. Surprisingly, the researchers concluded that people are separated by an average of six people. Communication and connectivity remain key in our perceptions of the world around us. Networks have enabled scientists to collaborate on important projects and discoveries that have changed our lives. New innovations have also provided us tools to detect hidden biological networks of Mother Nature. We often underestimate the importance of the underlying “water-cooler” conversations and unexpected encounters that lead to coincidences of ideas. Yet, as the world is constantly evolving, so is the process of networking. The Internet has allowed us to expand our networks like never before through sites such as Twitter and LinkedIn. Ultimately, networking remains a uniquely human characteristic that has given us a unified sense of direction and progress towards our goals.

As a reminder of the networks that surround our lives, this issue of the DUJS features the multitude of connections we often brush past and take for granted. Hae-Lin Cho explores the positives and pitfalls of social network sites. Ali Siddiqui describes the network phenomena behind climate change and weather patterns. Jessica Barfield details the ongoing conflict with the Internet and net neutrality. James Howe explains how our genetic importance rapidly loses its clout in only a few generations. Annie Sun describes how plants communicate using other organisms, namely fungi. Nancy Wu provides an overview of an emerging trend of massive open online courses. Jingya Qiu shows how molecular and pharmacological networks may be useful for applications in medicine, and Julia Isaacson explores the similarities in the neural networks between anorexia nervosa and autism. This issue’s two faculty interviews features Tillman Gerngross, Ph.D, a Professor of Engineering at the Thayer School of Engineering at Dartmouth, and Lee Hachadoorian, Ph.D, an Assistant Professor of Geography at Dartmouth College.

This issue of the DUJS also features three original research submissions from Dartmouth students. Matthew Jin reviews activity in Clean Development Mechanism networks that reveals fraudulent behavior. Alex Gerstein and Scott Gladston present findings that shed new light on file sharing behavior of Dartmouth students and provide policy recommendations based on their data. Finally, Chase Schoelkopf looks into the effects of habitat disturbance on parasites inhabiting the forests of Madagascar.

This issue of the DUJS would not be possible without the diligence and dedication of the writers, editors, staff members, and faculty advisors. We would like to thank all of those involved in maintaining the journal’s continued success as an outstanding undergraduate scientific outlet.

Thank you for reading the DUJS, and we hope you enjoy this issue.

Sincerely,

Steven X. Chen
Editor-in-Chief

and

The DUJS Editorial Board
# Table of Contents

**Faculty Spotlight: Tillman Gerngros** ................................................................. 3  
Shinri Kamei ’16

**Faculty Spotlight: Lee Hachadoorian** ................................................................. 7  
Arunsrinivasan Ponshunmugam ’17

**Social Networking: A Double-Edged Sword** ......................................................... 9  
Hae-Lin Cho ’16

**Concerns with a Heating Nation** ....................................................................... 12  
Ali Siddiqui ’17

**The Internet and Net Neutrality** .......................................................................... 16  
Jessica Barfield ’16

**Distant Cousins** .................................................................................................. 19

**Why We Are All Royals, and Why It Means Absolutely Nothing** ...................... 22  
James Howe ’17

**Underground Networks: Plant Communication via Fungi** .............................. 22  
Yi (Annie) Sun ’16

**Massive Open Online Courses:** .......................................................................... 25

**Bridging Technology and Education** .................................................................. 25

**Network Science:** .................................................................................................. 28

**The Future of Medicine?** .................................................................................... 32  
Jingya Qiu ’16

**Is Anorexia a Female Presentation of Autism?** ....................................................... 32  
Julia Isaacson ’15

**ORIGINAL RESEARCH**

**Structural Holes, Brokers, and Fraud in a Clean Development Mechanism (CDM) Market** ......................................................... 35  
Matthew Jin ’17

**Perceptions & Use of BitTorrent P2P File Sharing by Dartmouth College Students** ......................................................... 40  
Alex Gerstein ’15 and Scott Gladstone ’15

**Effects of Habitat Disturbance on the Parasites of *Propithecus edwardsi*** ........ 46  
A. Chase Schoelkopf ’15
Faculty Spotlight:

Tillman Gerngross

A researcher and entrepreneur, Thayer’s Professor Tillman Gerngross has never been afraid to hold back when push comes to shove. Despite being told that he would never succeed, he developed a groundbreaking method of making protein drugs from yeast and made his first mark in the entrepreneurial space by founding GlycoFi. Six years later, in June 2006, Gerngross sold GlycoFi to Merck & Co for $400 million, a deal that was “the third-highest price paid for a private biotechnology firm, and the largest on record to be done in cash,” as reported by The Boston Globe.

Since then, Gerngross has been named a 2013 National Academy of Inventors Fellow, an NH Entrepreneur of the Year, and more. But success has not made him complacent, and Gerngross has continued to innovate, founding Adimab, now valued at more than half a billion dollars, in addition to Arsanis, Avitide, and Alector. Read more about this serial entrepreneur’s experiences and what he has in the pipeline.

Could you tell me a little bit about your work and what has defined your path?

There are a lot of pieces to it. But one is a very Dartmouth-specific piece. I used to be in Cambridge; I was at MIT for many years in the biotech industry there, and what I really learned after coming to Dartmouth is this: when you’re in one of those major hubs, there’s always a ton of news flying around all the time. People are really excited, in a way, with what the buzz is at the moment. When you’re here, at least in my experience, you have to come up with your own story and your own rationale. You have to go out and find the data, interpret the data, figure out what it means, and come up with your own hypothesis, as opposed to being this little cork that floats around in the ocean with all of these influences hitting you from all directions. So for me, being isolated was an advantage, not a disadvantage. And it really allowed me to think really hard about the issues that I was paid to think about as a professor. It brought me to some of the conclusions that were unorthodox at the time but are now becoming mainstream. But, there was a time when most people thought: “this guy is crazy.” In a way, I guess you have to go through the phase when people think that. Unfortunately, there are some people that always stay crazy (laughs), but there are also people that end up being very successful in what they do, and then it sort of becomes legend.

Could you tell me about GlycoFi and Adimab and how those companies came to be?

The first job you do is the launch pad for the next thing you do. I had worked in a privately owned biotech firm in Cambridge, working on a particular problem. The problem we were working on was how you make a plastic from a renewable resource. In fact, you can make this type of plastic from microorganisms. Microorganisms can take sugar and make this plastic. And the view of the world at the time was: “that’s a wonderful thing,” so at this company, we were trying to figure out how to feed a microorganism sugar to make it into plastic, and everyone thought that was the greatest thing ever because the plastic was
made from a renewable resource, which sugar arguably is, and it’s biodegradable, which many people think is a good thing. The motivation to work in this area is that it’s environmentally friendly, but nobody had actually looked at it. Nobody had figured out what it takes to grow the corn, or what it takes to get the sugar out, because there’s a lot of fossil fuel input. People have said that corn is a crop that is drenched in oil. In order to fertilize it, irrigate it, and harvest it, you have a lot of energy inputs that are actually derived from mostly carbon. There’s a certain amount of fossil carbon that has to be expended to make each pound of corn. So if you then think about making stuff out of that corn, you can trace it all the way back and say that in order to grow the corn, to process it to make this, how much fossil carbon do I really need? And it turns out that when you do the math, this (holds up plastic container made of microorganisms) takes more fossil carbon than this plastic.

In other words, this isn’t really a solution that is more environmentally friendly. It’s a really hard thing to understand. I’d invested seven years of my life—two years at MIT and another five years at this company—in developing this process, and when I realized then that it wasn’t a good idea wasn’t easy. Very few people are willing to confront themselves with the reality of just having wasted seven years of their life. I learned a lot. So it’s not that I really wasted it, but it was not an objective that made any sense to me at that time. So here I am; I just showed up here. Within a year, I did all the math and figured this out, found that this process makes zero sense, I published it, and again, people said, “That guy is crazy.” I was among hundreds of other scientists working in this area across the world, and they hated me with a passion. I was really made fun of. I was among hundreds of other scientists working in this area across the world, and they hated me with a passion. I was really wasted it, but it was not an objective that made any sense to me at that time. So here I am; I just showed up here. Within a year, I did all the math and figured this out, found that this process makes zero sense, I published it, and again, people said, “That guy is crazy.” I was among hundreds of other scientists working in this area across the world, and they hated me with a passion. I was basically viewed as someone who was spoiling their efforts in a way. But it is what it is. I’m not going to spend the rest of my life doing something that makes no sense.

It sensitized me to one thing. That is, once you think about solving a problem, think very hard about what is the quality of the problem you’re solving. This is something that scientists are so poorly trained at. We are like greyhounds: if it’s fuzzy and moves, it has to be a rabbit. It turns out not so, right? That’s how scientists deal with problems. They immediately see a problem and they want to solve it, but they don’t think of the millions of problems that they could be working on, and which ones are more or less important. That was my big takeaway from this whole process.

The human genome was about to be completed. There was an expectation that this was going to lead to a bunch of drugs. How are we going to make those drugs? At that time, people were talking about this bottleneck of how we were going to make those drugs based on protein therapeutics. I had an idea of doing it differently. People said that’s some wonderful stuff, and I wrote grants, like every academic does: here’s how I’m going to solve the problem, sent them in to the NIH and all the other funding agencies. And they said listen buddy, this is a really hard problem. You have no experience, no specific expertise in glycobiology, which is what’s required. They said no funding for you, and, by the way, if someone’s going to solve it, it’s not going to be you. (Laughs). As an entrepreneur, you have to get used to the fact that people are not as enchanted with what you are trying to do as you are. It doesn’t mean that you’re not going to be successful. In my case, they were very specific about why they thought it wasn’t going to work out. To make a long story short, it worked out.

But where did you get your funding?

Like I said, I wrote a lot of grants, but no one wanted to fund the company. And then came the network of people here at Dartmouth, in particular at Thayer, who have been out in the world. At Thayer, we have something called the board of overseers, and it’s really a tremendous asset. The board of overseers is either Thayer alums or people affiliated with the school who have gone out into the world. The real value that they bring is a view of the world that is not just academically influenced. So when I talk to these guys, this is really interesting. When I was teaching a course, the former dean of the engineering school, Charles Hutchinson (waves at picture behind his desk), that guy over there to the right in that photo, so he said, that sounds really interesting. Do you see those napkins underneath? (Waves at framed napkins under photo of himself and Hutchinson). Those are the original napkins that we used over at Tuck (Byrne Hall) to map out how we were going to start the company and what we needed to do. He gave it to me as a gift when we sold the company in 2006.

He and I used the Dartmouth network to connect to two people. One was a Thayer overseer by the name of Michael Ross who had been in the biotech industry for 20 years and was very knowledgeable in that industry. The other one was a guy named Terry McGuire, who is now the chairman of the board of overseers. And Mike said: “This is a great idea. If you can crack this nut – it’s hard, but if you can – it’s very valuable.” Terry was like okay, that’s exactly what I want. Let’s give these guys a little bit of money and see what they can do with it. So he funded the company. He and Angela Parvi funded the company with a few hundred thousand dollars and within a year, we could do more with that money than anyone else had ever done before and they saw this and said that this is an effort worth backing. That’s how we raised more money and, ultimately, the company raised about $30 million in total. They saw that the scientific success was there. We published in Science and Nature and Biotechnology, among others, and companies were coming to us, asking us to help them solve their problems. And when they see that, they realize that these guys are working on something that’s really going to solve someone’s problems, and sure enough, the company was acquired in 2006. And they made out really, really well. And they all had huge returns on their investments, and everyone was a happy camper.

Could you tell us how well?

The early investors got thirteen to fourteen times their money back, which doesn’t happen very often, unless you buy lottery tickets or something.

And how much did the company sell for?

$400 million. And everyone thinks that I got all of that
I read that you were unhappy with the tech transfer program that was in place when you were originally developing your technologies.

And now you’re the provost of the entrepreneurship and tech transfer office!

So that came about because, unfortunately, there were various people working in that area, but they weren’t working very well with one another, and there was a lot of unnecessary friction. At some point under President [Jim Yong] Kim, then-provost Carol Folt asked if I would do this, and I said, “yes, but here are the rules. I want a budget for a certain number of years, and I want to hire people that can do this day-to-day,” and that’s when we hired Trip Davis and Jamie Coughlin. So those are the guys now that are really going to provide now and set up the infrastructure of a sustainable presence supporting innovative students and innovative faculty. We are really excited to help. The alumni are behind this, and with the Innovation Center (scheduled to be built at the start of next year), there will be all sorts of programs running, and it’s going to be very exciting. The problem that I often find is that I meet students, but my schedule is so packed that it’s very hard for me. If I met with every student on this campus that had a startup and I could basically throw my day job away. We want to make sure that we have programmatic opportunities, that we’re in the innovation program, then I will be there and have opportunities to interact with the students, which is just not possible right now.

“Great academic institutions stay great academic institutions or become better by attracting the best people. And we do a phenomenal job at the undergraduate level. I think that our undergrads are way above the average. And I look at our faculty, and I think that we’re above average. But I think that we can do better.”

It’s worked out well for me. But this is mostly my thought process. Great academic institutions stay great academic institutions or become better by attracting the best people. And we do a phenomenal job at the undergraduate level. I think that our undergrads are way above the average. And I look at our faculty, and I think that we’re above average. But I think that we can do better. I think that the problem with academia when it comes to assembling faculty is that your currencies are very limited. For the super-talented molecular geneticist that has an offer between us and Stanford, where’s she going to go? There’s a high incentive to go to Stanford because of name recognition around research, and if she has a spouse, that spouse is going to have to find a job here. Smart people tend to marry smart people, so that spouse is going to be a lawyer or something, so if they can’t do anything, they’re not going to come here. So my view is that we need to change that and find ways of encouraging people to come here and give them something that they can’t get anywhere else. And salary’s not it. You can’t buy yourself out of this problem. You have to figure out certain things that you have to be really good at and that you can really support. For me, one example is the whole area of tech transfer and how we think about enabling our faculty to do great things in the world. I’m one person here. And at this point, there are one, two, three, four, five molecules that have gone into humans for testing to become drugs that could change people’s health. That’s the ultimate success, right? Forget papers and all this publishing that people do. It’s important, I’m not saying that it’s not. But at the end of the day, we need to find metrics that really matter. And impacting human lives, I think, is a very noble thing to strive to change for the better. So, what I’m saying is, why don’t we think about it from that perspective? Why don’t we help our faculty impact human lives and get the tech transfer office out of the business of taxing innovators so that the college owns a little piece of it, which is how most tech transfer offices operate. They are taxers of innovation.

Has that changed recently with your work?

First of all, they put me in charge, and, believe me, they wouldn’t have picked me if they’d wanted someone that does it the old way. They picked me because they thought that I could bring about change, but change in academia can take time. I don’t have the power to walk in and say, guys, this is how it’s going to be. You have to go through the committee; at two o’clock today, I have a committee, a task force for four hours, to discuss the changes. And this is probably meeting number 10. There are a lot of issues to be discussed, but I’m willing to put in the time and effort to change things. Once we change the policy, then you can go to the world and say, hey guys, here at Dartmouth we do things differently. One of the things that bugged me when I first came here was that people are always saying: “What are our colleagues at Harvard doing?” I say, screw that. We’ll figure it out on our own. If we just follow
them, we’re never going to out-compete them. We’re going to have to focus on things that are true to our principles and our values and what we care about. We have unique consolations here. We have a business school right across the street here (gestures at Tuck, right outside his office window). Literally.

There’s not even a street here!

Exactly! I love that! There’s not even a street here. We have to build upon our strengths. Focus on innovation and entrepreneurship with a business school not even across the street. Focus on building an alumni base. If you look at the venture capital community, there are people that spend their entire careers figuring out how the world is and how the world could be better, and here are people that could make that change happen. If I put money behind them, that change is going happen, and I can make money in the process. That industry is called the venture capital industry. That’s all they do. Figuring out all the opportunities to make the world a better place. Perhaps I’m painting a rosy picture here because some of them just want to make money, but, by and large, it’s finding opportunity and finding out where the opportunity of capital plus talent leads to change and creates value. There’s no industry where Dartmouth alums are more represented than the venture capital industry. I don’t know why, but it is a fact. Go to any big VC firm in the country, and there are going to be alumni there exquisitely trained in exactly this thing. We have a business school across the street. We have another strength: we have a strong presence in biochemistry, biology, and all that stuff, and we have a medical school here. What does the world of drug discovery look like? People figuring out how stuff works, and people, based on that, turning those discoveries into drugs that then end up benefitting hospitals. There are doctors and basic scientists involved in that process. We have a unique opportunity, so I say: great, let’s leverage that and build upon it. I feel that we should build upon what we have.

What’s fun is that it allows you to continue to be engaged with students, and it’s all about fun at the end of the day. I’m in a fortunate position where life is about maximizing fun and giving back to this institution that gave a lot to me. It’s a lovely position to be in.

And what are those? (Pointing to a row of champagne bottles of increasing size on top of the cabinet)

Because... so you see those napkins over there? My cofounder, Hutch, said: “This science is way too complicated. No one is going to understand this. We have to break this down in very discrete milestones.” So we came up with six milestones. The first milestone is that little bottle—so we celebrated our first milestone with that. And then as the milestones progressed, we got a bigger and bigger bottle.

Could you explain your milestones?

The scientific problem in GlycoFi that we tried to solve is engineering a yeast strain to make human proteins that have human sugars on them. So it was a very complicated cell engineering problem.

How long did it take you?

It took us six years. (Chuckling). The problem is, the bottles are getting so big now...
To you, what is a network?

Oh wow, we are going to start really philosophical, huh? A network is a series of relationships among entities. A social network is a series of relationships among people; a street network is a relationship among places on a street.

What is Geography?

Geography is a very diverse field that includes people doing physical science including climate science, geomorphology, and coastal processes. It also includes people researching social sciences, looking at population movement, urban form, etc. Even within that side of Geography, there are people that approach the subject more quantitatively. Others pursue the field more qualitatively and theoretically. It is an extremely diverse field, usually unified around the idea of studying processes on the habitable portions of the Earth, or the parts of the Earth that impact humans or are influenced by humans. There are a number of themes throughout the history of geography including the relationship between culture and the environment or humans and the environment. Geography includes spatial analysis and looking at how processes unfold in space and develop spatially. There are a variety of themes that tie Geography together – it is hard to sum it up in a statement.

What is the public perception of Geography?

They usually have no idea what the hell you are talking about. The public perception in this country is invisible, nonexistent. It is still taught as an area of study in a lot of other countries, particularly in the Commonwealth countries. So, if I have students coming from the UK or India, along with many other countries, they will be familiar with geography as a discipline. Very frequently, a student comes from outside the United States and has been exposed to the idea of geography and knows what it is. To the extent that people in this country are aware of what geography is, they probably think that it is memorizing a list of capitals. They are not familiar with the idea of geography as a science, what questions geographers ask, and how geographers try to answer those questions.

How would you describe the Geography department at Dartmouth?

This is not a surprise to anyone, but the Geography department here is very socially and culturally oriented. Most of the faculty here do not do the same kind of research I do. Xun Shi is the main Geographic Information System (GIS) professor here; Jonathan Chipman, who runs the lab and teaches environmental GIS, is also a GIS person. There are a couple of physical geographers and most of the department does human geography. Within the part that does human geography, a lot of them are doing human geography from a social and cultural perspective. That is the strength here in terms of the research focus. Beyond that, it’s a very collegial department. We are pretty small. In general, the faculty here are very friendly. There is a lot of collaboration. A lot of the faculty put out grants with each other and work on projects with each other, so I would say that it is a very collegial department.

What are GIS and spatial analysis?

Geographic Information System or Geographic Information Science refers to the use of computational resources for the storage, analysis, retrieval, and display of geographic data. Anything that exists in space can be geographic data. This includes physical and human processes. It could be used for mapping trees in a forest; areal aggregates like the population in a town, or some economic characteristic like the poverty rate. You have the world, and GIS allows you to model the world into a form that can be stored on the computer. In town governments, for example, GIS could act as a simple retrieval system, and it can map networks of streets and networks of where all of the fire hydrants are. Beyond that, it can be used for display in order to create informative and attractive maps of different processes. It can also be used for analysis. I study residential location, and I am interested in why people live where they live, so I use the GIS to store various economic and demographic data – data about where local governments are and characteristics of local governments. I also use GIS for analysis like how much the population is economically or racially segregated.

How are networks applied in each of the various subsets of Geography?

If you are looking at population Geography and looking at why people move where they do, then chain migration networks come into play – people move in paths that others go down. Within GIS, which may get used for asset-management or transportation analysis, physical networks may be modeled directly in your GIS, trying to figure out the least cost path from point A to point B, which may involve moving across rail networks and street networks. So that physical network may be directly modeled in the GIS. Also, we are moving into the study of emerging social networks and how people influence each other through social networks. Since human networks exist within the space, there is a question of how social networks are created and influenced by the spaces in which they are created.
What does Geography research look like?

It really depends on your subfield. Researchers in qualitative geography frequently are involved in some kind of fieldwork that actually involves interacting with subjects to gather their data. In physical geography, it depends again on the subfield. If you are doing large scale climate modeling, you may largely rely on sense data, such as satellite data and precipitation measurements. In this case, there is a lot of computation and may not involve a large field component. On the other hand, conducting watersheds or stream analysis may require fieldwork from stream gauges or investigating particular areas to see how accurate measurements are. It varies widely. My own research tends to be very much computer based and quantitative. Often I am relying upon open, publicly accessible data sets like the census, government economic data, or social data from cities. I have not done my own data gathering so my work is very computer based. I am interested in taking data, turning it into variables, developing a model, and either mapping it or performing statistical analysis to look for relationships.

What does your research look like?

I am mostly interested in urban areas and urban spatial forms. Within that I am interested in where people live, why they live there, the history of suburbanization, and the current inversion where the population, to some extent, is returning to the cities. My research involves how this ties into urban policy, inequality, and poverty. My overarching goal is to see what makes cities work or fail to work and investigate ways in which they can be improved. So it broadly falls within urban geography and population geography because a lot of my research focuses on demographic and economic data. I have a paper I am revising right now that looks at the extent to which people are separated into homogeneous governmental units in suburban areas. I looked within New York, comparing Nassau County and Queens County to see whether or not there are sharp divisions at government borders. We know that there is a fair degree of clustering of population along demographic and economic lines in cities. The question is: do local governments facilitate sorting into homogeneous communities? What I found is that they do. The question then becomes: what do you do about that? Within the public choice literature, this is actually a good thing because it allows people to choose local governments that match local governments that match their demand for particular services. On the other side, there is a concern that this leads to socioeconomic and racial segregation, which makes it harder for the disadvantaged to improve their economic standing. What is interesting about this from my point of view is that there are strong arguments from both sides. That is an area of my interest in relation to urban research.

Some of my research is more technical. One of the things I am trying to get started is dasymetric mapping, which is a method for estimating population. There is a population reported at the census track level and you don’t know where in the census track the population resides – it could be clumped on one side or the other. In dasymetric mapping, you use ancillary data like where are their water bodies, where are their parks, where are their buildings, and where are there taller buildings versus where are there smaller buildings. This information is used to figure out where to distribute the population within that aggregate. These methods have around for a while but I have been working on right now is to develop some computational tools that would automate the process; otherwise, this is a laborious process in spatial analysis that requires expertise. I am trying to develop tools that would make it easier for non-experts to use this kind of analysis.

What are your favorite classes to teach and why?

That is a tough one. Because of my particular interests and my area of study, I tend to teach both technical classes as well as topical, urban focused classes that do not have a technical requirement. I really enjoy both of them; I like having that mix. The topical classes get into the issues that, I think, are exciting. The technical classes are just fun because I like playing with the technology and seeing other people use it. The students often use the technology for projects that I would never come up with because that is not my area of interest. My favorite class so far must be a class I taught last term [Winter 2014] for the first time in Geovisualization. It was the first time Dartmouth offered it. It was incorporating aspects of computer cartography as well as data exploration techniques. I was teaching students how to make a flow map, how to visualize multiple variables at the same time, as well as web based cartography. I really enjoyed it. Besides the topic, the message I was trying to get across to students was, in the real world, you will face problems, perhaps in a job or a project you are interested in, and you need to rapidly identify what tools you are going to use, get up to speed on them, produce results and document them. A lot of the class was focused on learning how to learn. In addition to teaching students particular methods of various software, they were given a topic, like how to use a flow map, and they would identify the software and build a full tutorial on how to use the software. I think that they really enjoyed it and got a lot out of it. It was a very hands-on, interactive class – a class where I learned a lot from the nature of these assignments and they learned a lot from each other.

Do you think more classes should be taught in a more interactive style?

Yes. There is very little that is less interesting than standing up in front of a room and lecturing nonstop for an hour and five minutes, three times a week. So, as much as possible, you try to get the students to become their own teachers and their own guides. I don’t know if I succeeded in that goal in every class but at least in my Geovisualization class, I think I did get a lot of the way there.

How important is networking in the sciences and social sciences?

I think pursuing research by yourself is a possibility but often, people are looking for someone else with similar interests to research with. So, networking is important in terms of finding someone who has similar interests and trying to figure out if there are projects to work on together. These days, many projects, particularly in the quantitative and the physical sciences side, are multi-investigator projects that bring in people with different sets of skills and different expertise but, similar interests to conduct research. That is one of the more important parts of networking in the sciences. In terms of finding jobs, it depends on whether the job is inside academia or outside academia. I think that outside academia, networking is tremendously important for finding jobs. Inside academia, I would not say that it is unimportant, but there is a lot more emphasis on credentials and scholarships. You are committing to years with a candidate when making an academic hire. Hiring someone into a professorship, if does not work out, is kind of a problem. Hiring in academia often involves ensuring that you have a budget line that will pay a person’s salary. Getting the wrong person sets you behind for years to come.
Introduction

In the 1960s, Stanley Milgram asked 296 people to deliver postcards to a stockholder from Boston, with a catch – in a game of probability, participants could only send the postcard to a personal acquaintance who was expected to be closer to this “target person.” What came to be known as the “small world experiment” determined that there were between 4.4 and 5.7 “degrees of separation” between the receiver and sender of the postcard. In other words, on average, any two people in the world were separated by roughly 5 people (1).

In the 21st century, the average degrees of separation between Facebook users has fallen to just 3.74 (1). Our world, in the face of modern technology, is rapidly shrinking.

Social network sites (SNSs) like Facebook aim to “connect” people to each other. The word connect is featured in the mission statements of Facebook and several other popular SNSs like Twitter, LinkedIn, and Instagram. Recent statistics show that 65 percent of internet-using adults in the US use SNSs and that about 90 percent of undergraduate students use Facebook, which passed one billion users in October 2013 (2,3).

In spite of their growing popularity and influence, people have started to question several aspects of SNSs. Although these networking sites are supposed to help people build, maintain, and strengthen relationships, recent developments suggest that the psychological repercussions of these sites may leave users feeling worse than ever.

What are SNSs?

Nicole B. Ellison and Danah Boyd defined social network sites in The Oxford Handbook of Internet Studies:

“A social network site is a networked communication platform in which participants 1) have uniquely identifiable profiles that consist of user-supplied content, content provided by other users, and/or system-provided data; 2) can publicly articulate connections that can be viewed and traversed by others; and 3) can consume, produce, and/or interact with streams of user-generated content provided by their connections on the site (4).”

The definition, revised from an earlier 2007 version, highlights the more interactive and interlinked nature of present-day SNSs. With more developed technology, sites have started to offer additional features that allow for faster, easier daily updates, comments, and uploads. The definition of a “friend” has also increased in complexity with the addition of features such as following, hiding, and creating social circles (4).

Despite the difficulties of studying the many facets of SNSs, which continue to grow and evolve, many researchers have focused on specific angles like privacy and communication, and have placed an increased emphasis on the effects of SNSs on the cognitive and emotional states of their users (4,5).

Considering SNS Users

In the face of Facebook’s skyrocketing popularity, experimenters have targeted the social network site to study whether it can increase the quantity or quality, otherwise known as the “social capital,” of interpersonal relationships for users.

Currently, two opposing theories exist on this topic: the poor-get-richer and the rich-get-richer. According to the first theory, people who lack the social skills to build relationships turn to social network sites for online support. From this perspective, sites like Facebook and Twitter enrich the social capital of those with few friends in the real world. On the other hand, the rich-get-richer theory claims that social network sites do not necessarily benefit
people who do not have friends offline. In fact, these more introverted users may have fewer friends online than their socially skilled counterparts, despite spending more time on Facebook. Current evidence suggests that the latter theory fits better with observed trends. It is generally accepted that using Facebook as a replacement for real world interpersonal relationships is not beneficial to the social or emotional health of the user (2).

Considering the SNS itself

Early on, the bulk of scientific research on social network sites considered how the personality type or emotional state of users might affect the way in which that individual would be affected by online social networking. However, recent findings suggest that, at least for Facebook, the social networking site may prove deleterious for all users, rather than just those who were already socially isolated (5).

Scientists at the University of Michigan found that Facebook users suffered a decrease in subjective well-being, a multi-faceted measure that also predicts for future health and lifespan. To test the short-term effects of Facebook, participants in the study responded to an online survey five times a day that asked them to rate their current affective state, level of worry, and level of loneliness. The more often they used Facebook, the more likely they were to feel worse later on in the day. In addition, at the end of the two-week long study, participants experienced a decrease in life satisfaction (5).

While results revealed that people who felt lonely were more likely to use Facebook, the findings remained significant even when the researchers controlled for loneliness and baseline emotional states. Several other variables, including number of Facebook friends, motivation, and depressive symptoms, also showed no effect (5).

Scientists do not yet fully understand the mechanisms behind the decrease in subjective well-being associated with the use of SNSs, but Festinger’s social comparison theory grounds one prominent theory. According to Festinger’s social comparison theory, people evaluate themselves by comparing to others, especially in ambiguous situations (3). Facebook and other social network sites provide the perfect platform for people to make comparisons, having access to details about numerous other people in their social circles (3,5,6).

Despite being used for networking, SNSs also require “self-presentation” (4). Due to the nature of online profiles, SNS users can choose to edit and refine which parts of their lives they want to show to other people, which can give false impressions of how others are living (6). In fact, the more times people used Facebook, the more likely they were to make social comparisons to others (3). In addition, the frequency of social comparisons made on Facebook was positively correlated with negative emotions from these social comparisons (3).

These effects were most evident when Facebook users passively interacted on the SNS. The amount of passive following was positively correlated to feelings of jealousy, which were also negatively correlated to life satisfaction. This sets off a vicious self-promotion - envy spiral that results in users elaborating on the positive aspects of their profiles to match those of others (6).

Social Media Addiction

Although the simple solution seems to be to reduce time spent on social network sites, the addictive nature of social network sites, especially for young adolescents, complicates the issue (7,8). According to Professor Sherry Turkle, author of the meta-analysis on social technology Alone Together, the younger generation has become accustomed to a culture of social network sites, and they often fail to build the skills to manage social relationships without the technology (7).

“It’s a great psychological truth that if we don’t teach our children how to be alone, they will always be lonely,” Turkle said. “When they’re always connected, children, adolescents and adults become dependent on the presence of others for validation in the most basic ways” (7).

When separated from social technology, adolescents become disconnected from necessary social interactions, leading to negative effects (2,7). While research indicates that reduction of time spent on Facebook reduced negative side-effects like distractibility and negative emotions, it also decreased the positive emotions usually produced by active online social interaction (2).

Like drugs, the addictive nature of SNSs is related to dopamine, a neurotransmitter involved in reward and conditioning. Humans demonstrate a desire to boost self-esteem and a fundamental attribution bias. This tendency towards self-bias parallels recent findings that humans find “self-disclosure” inherently rewarding. Focusing attention on oneself, especially when sharing information with others about the self, activates the same brain regions associated with primary and secondary rewards in Pavlov’s classical conditioning: the nucleus accumbens (NAcc) and the ventral tegmental area (VTA) of the mesolimbic

“People who felt more lonely were more likely to use Facebook.”
dopamine system. When given a choice, people even chose to forego earning more money in favor of talking about themselves more (8).

As previously discussed, SNSs like Facebook provide a place for “self-presentation.” These sites rely on people’s active and willing participation in sharing aspects of their daily lives through updates, photographs, and videos (4). If this is the case, SNSs activate the mesolimbic dopamine system and reinforce the habitual behavior of logging in and talking about the self, which seems to be the conjugate of the self-promotion - envy spiral theory (6).

Conclusion

At the moment, no one really knows the extent of the summative effects of SNS use. Evidence suggests that a combination of multiple variables, including inherent personality, emotional state, self-esteem, environment, and use of SNSs could impact the psychological effects seen in SNS users. However, what is clear are the profound links between SNS use and a negative impact on mental health. It is unrealistic to try to purge the world of SNSs, which possess an enormous hold on the younger generation and come with a host of benefits, including faster communication of knowledge to a large audience, a platform for social activity and reform, and effective marketing strategies for corporations (4).

In the face of such an evident problem that is harming the mentality of a significant portion of the population, steps must be taken to alleviate some of the psychological effects. Some people believe the best way is to practice balance and moderation in all technology use (2). Others believe researching the strong ties to low self-esteem and SNS use may elucidate matters (3).

CONTACT HAE-LIN CHO AT HAE-LIN.CH0.17@DARTMOUTH.EDU

References

The persistent drought in California over the past year has sparked concerns about the effects of climate change on extreme weather events. Droughts have had a historic presence across much of North America. However, some weather and climate data from the Midwest drought of 2012 suggests increased risks of drought in the future. These additional risks may be due to the cycles of La Niña in the Pacific, changes in Arctic wind patterns, and increased average temperatures as a result of climate change. With the additional risk for drought comes the possibility that droughts like the one in California may last much longer than expected.

California suffered a drought in 2013, with many major cities like San Francisco, Los Angeles, and Sacramento posting records for low annual precipitation (1). California gets much of its annual precipitation in the winter, but the winter this year has not offered much relief after the drought of 2013 (2). A persistent high-pressure air mass off of the coast of Canada, dubbed the “ridiculously resilient ridge,” may be the cause. The system has been in place since January 2013 and may be changing the flow of the jet stream in California. A jet stream is a fast stream of winds that results from the meeting of large cold air masses, like those of the Arctic, with warmer air masses. Due to the Earth’s rotation, the jet stream near California typically goes west to east—from the Pacific toward California and the rest of the United States (3). The ridge has diverted the jet stream to the north, thereby directing many storms that are pushed by the jet stream from the Pacific toward Alaska instead of the West Coast (4). Blocking ridges are a normal phenomenon, but they typically only last for a couple of months (4). This particular ridge has lasted over a year – long enough to have changed the temperature distributions of the region’s air as well as local waters, thereby making it more difficult to break up the ridge (4,5). Continued warm sea surface temperatures south of Alaska may be a result of this effect. The sea surface temperature could have played a role in what seems like a revival of the ridge after a few storms made it through to the parched coast (5).

As of April 2014, those storms have left precipitation levels at less than 50 percent of normal for much of the area. The system is still dynamic, and the ridge could disappear, but the incoming summer season could minimize any changes to reduce drought-like conditions (2). The unusual occurrence of a prolonged blocking ridge, paired with the already severe drought, has led many to suspect that climate change may be playing a role in the situation. Although the current drought is too recent to establish a strong relationship with climate change, some suspect that because climate change increases the odds of other extreme weather events, like heat waves and thunderstorms, climate change may make long-lasting blocking ridges more common. Although more research is needed and many of the effects of the current drought have yet to be seen, another recent drought, the Midwest drought of 2012, may provide insight into how patterns of climate change affect drought.
The 2012 drought reflected dry conditions across much of the nation, not just the Midwest. About 81 percent of the contiguous United States experienced atypical dryness (6). The drought-like conditions in the Midwest were of particular concern because the region is an agricultural center. Researchers noted that this drought occurred after a period of La Niña in the Pacific Ocean during 2010 and 2011. La Niña refers to times when the ocean temperatures of the tropical Pacific are cooler than average (7). During 2012, conditions in the tropical Pacific were still quite similar to those seen during La Niña, leading to speculation that the drought may be related to this lingering effect. Such speculation is warranted, for there is a history of correlation between La Niña and drought, at least in the American Southwest. Often, when La Niña is present in the Pacific, reduced precipitation and drier conditions in parts of the United States is observed (8).

The La Niña-drought trend has been observed throughout the 19th and 20th centuries. During the 19th century, La Niña was observed for periods of up to nine years, but in the 20th century, La Niña never lasted for longer than two years when it appeared in the Pacific (9). Regardless of the length of La Niña, droughts in the Southwest corresponding to La Niña have been observed to expand into the Midwest, sometimes even after La Niña has disappeared in the Pacific. Such instances show that the La Niña-drought correlation is not perfect. La Niña cycles seem to explain 30 to 50 percent of the variance in drought occurrence. Not all La Niña events lead to drought, nor are all droughts caused by La Niña. Droughts also vary in severity and in the areas they cover. For example, a drought that corresponded to La Niña from 1859 to 1862 was originally localized to parts of Texas and New Mexico, but the drought spread through much of the West from 1863 to 1865, after the Pacific had returned to normal and warmer conditions than those of La Niña (10). Multi-year La Niña events, like those of the 19th century, have not been observed in recent years. Historic multiyear droughts are rare but regular, and have been observed “about twice per century in the past 400 years” (10). These kinds of droughts include those that caused the Dust Bowl in the 1930’s, a period during the Great Depression where conditions were dry enough that the loose soil in the Midwest could be blown into dust storms (10).

The conditions that exist during these droughts can be worrisome. As seen with the blocking ridge in California, long-term weather conditions have the potential to alter local temperature and climate so as to prolong abnormal conditions. In the case of any drought, reduced moisture leads to less evapotranspiration, which further decreases the water available in the area (10). Local temperatures can then rise and lead to even drier conditions (6). The worry today is that the increases in temperature due to global warming may make a drought more likely to become a multiyear drought because local temperatures are already elevated. As in the past, it may also be possible for droughts to outlast or to form after the end of La Niña due to predisposing factors like elevated temperatures.

Aside from drought, increased temperatures due to climate change can also lead to increased evapotranspiration, which dries the plants and the soil, thereby continuing the cycle of water loss and rising temperatures typically observed during drought-like conditions (11). Another effect of elevated temperatures is the melting of snowpacks. Warmer temperatures in the late winter or early spring can hinder the development of snowpacks and cause them to melt too early, thereby reducing water availability later in the year and elevating the risk for drought (6).

Besides La Niña and elevated temperatures due to climate change, changes in Arctic wind patterns may have affected the Midwest in 2012. Arctic wind patterns affect how precipitation reaches the United States. During the summer, these winds normally flow from the west to the east. However, from 2007 to 2012, they have had an increased tendency to move in a south-to-north pattern. Aside from the fact that these changes tend to accelerate the melting of ice sheets due to hotter summer air from the south coming into the Arctic more frequently, such changes can also disrupt the flow of the

Figure 2: Increased risks for drought raise the possibility of multiyear droughts, like those of the Dust Bowl.
jet stream. The jet stream typically moves across the United States from the west to the east, following the flow of the Arctic winds. Recent changes in the Arctic winds may add north-south motion to the path of the jet stream (12). This additional motion slows the jet stream and the storms associated with it since the jet stream takes longer to get from the Pacific to eastern parts of the United States (13).

A slowing of the jet stream would lead to a long-term increase in the risk for extreme weather conditions, like heat waves, cold spells, flooding, and droughts (14). Without the jet stream bringing in new storms and weather conditions at a steady pace, new blocking patterns can form once a particular system has remained in place for some time. A heat wave in March of 2012 that affected the Midwest and Northeast may have been an example of this. The jet stream was stuck in Canada at the time, allowing for the warm air to remain in the Midwest for longer than usual (13).

So much of climate change is interrelated that it is hard to accuse one factor, like Arctic wind patterns, for an event like a heat wave. Many factors could work together. For example, La Niña seems to be related to the risk of heat waves as well. A study revealed that heat waves were 20 times more likely in Texas during 2008 than in those years affected by La Niña in the 1960s (15). Why the risk was elevated in 2008 relative to previous La Niña years is unknown. While the study cited elevated temperatures as a possible culprit, it could very well be that the changing Arctic wind patterns were playing a role as well.

Despite the uncertainties in cause, it would be wise for officials, industries, and the public to look at the increased risks for trouble and foresee the consequences of future droughts in the Midwest. For example, a large drought in 1988 cost more than $40 billion in losses (6). Another long-term drought today would be just as costly, if not more so. During the 2012 drought, 60 percent of farms in the United States faced drought conditions (16). These conditions led to decreases in production of many grains. Corn production decreased by 27.5%, and soybean production decreased by 7 percent (17). The United States is the largest exporter of both corn and soybean, so global prices increased due to the reductions in production seen in 2012. Corn prices went up 25%, and soybean prices went up 17%. These increases may also have played a role in the sharp 10% increase in global food prices seen between June and July of 2012.

By the end of the year, both corn and soybean prices had stabilized but remained higher than the prices seen before the drought that year. Often, these price increases can carry over for a couple of years (17). However, due to a record corn harvest and the third-largest recorded soybean harvest in the United States, prices in 2013 were able to drop to more typical prices (18). However, one cannot expect record harvests to always negate a particularly bad season. If the risk for drought increases, so too should the risk for elevated food prices in the long term. Elevated prices would impact not only the United States but also the consumers and economies abroad who depend on many of our crops.

All droughts have many factors that affect their probability and severity, and they also have many of their own effects. Although it is hard to pinpoint the current California drought...
on solely climate change and the blocking ridge, the drought’s presence and its associated consequences are unquestionable. Droughts lead to reduced precipitation and increased temperatures, two elements necessary for wildfires. In recent years, it seems that wildfires have been getting bigger. In Southern California, this pattern has been seen despite the fact that the number of triggers has remained the same. In this region, most fires are started by humans, either on purpose or by accident (19). Lightning is also known to cause wildfires (20).

Precipitation and temperature are not the only factors involved. Invasive species, like cheatgrass, can increase the risk of fire by adding to the dry fuel on the ground (19). In areas like the Northern Rocky forests, earlier snowmelt in the spring can elevate fire risks. These snowmelts may be a result of widespread temperature increases associated with climate change (21). Droughts could also supply additional temperature increases as well as increased evapotranspiration. Coupled with less rainfall, droughts may imperil the region’s water supplies in the future, affecting agriculture, yields, and food prices (21, 22).

Looking to the immediate future, this year’s wildfire season may be one to watch. So will any changes in the Pacific Ocean, where there are signs that an El Niño may form by the fall. El Niño is the warming of the waters of the tropical Pacific, and it may bring more precipitation to California. If it is a strong El Niño, enough precipitation may come to help relieve the drought. At the very least an El Niño would break up the blocking ridge off of the coast and allow for the jet stream and storms to make their way to California again (5). Even with short term relief, with long term risks for blocking ridges and droughts seemingly increasing, people must be prepared for scenarios where there may be prolonged periods of reduced precipitation or drought. The same would go for wildfires.

Many phenomena are at play in climate events. It would be convenient to pin the blame on future wildfires and droughts on only a single phenomenon – increased temperatures, La Niña, jet stream changes, and others – but they are all related in many different ways. As more data is collected, patterns and causations may become clearer than the correlations that researchers have found. In the meantime, the populations affected should be aware of the dangers associated with climate changes. It is also possible that the recent droughts happened by a coincidence of events. Nevertheless, seeing that the nation has had a history of prolonged droughts like the Dust Bowl, each drought should keep the nation aware that climate change may be playing a detrimental role in enhancing risk for extreme weather.

**References**

Introduction

Net neutrality is the idea that Internet service providers (ISPs) and governments should treat all data on the Internet equally and that Internet users should be able to access any web content and application they choose without restrictions or limitations imposed by their ISP (1). Proponents of net neutrality argue that broadband network providers be completely detached from what information is sent over their networks, and that no bit of information be prioritized over another. An assumption of net neutrality is that information networks such as the Internet are most efficient and useful to the public when they are less focused on a particular audience and instead more attentive to multiple users (2).

Under net neutrality, ISPs would not be allowed to prevent their customers from using competitors’ applications or to discriminate against competitors in more insidious ways like prioritizing their own video applications or allowing their Internet telephone services to work faster and better on their own networks than services offered by competitors (3). With net neutrality, a person shopping for a new consumer product online, would be able to shop on any and all websites, not just the website to which a provider has established a preferred business relationship. Lacking net neutrality, Internet service providers could discriminate or charge differentially by user, content, site, platform, application, type of attached equipment, and modes of communication (1,2). Many commentators believe that net neutrality is an especially important goal for the preservation of first amendment rights under the U.S. Constitution (4). They argue that net neutrality means an Internet that enables and protects free speech and does not attempt to inhibit speech or control the information a user is able to receive over networks (3,4). However, net neutrality is not a principle followed by all countries and all internet providers. Internet censorship in the People’s Republic of China (PRC), is conducted under a wide variety of laws and administrative regulations. In accordance with these laws, numerous Internet regulations to inhibit net neutrality have been made by the government of the PRC, which have been implemented by provincial branches of state-owned ISPs, companies, and organizations (see generally 5,6).

Impediments to Net Neutrality

Four aspects of the Internet marketplace and architecture are relevant for the debate on net neutrality. First, “backbone networks,” which consist of interconnected, long-haul fiber optic links and high speed routers, are capable of transmitting vast amounts of data. Internet users typically connect to these networks and ultimately to one another through local access providers like Verizon, who operate the “last-
mile” transmission lines. Second, access to the Internet is generally furnished through “broadband,” i.e., high-speed communications technologies, such as cable modem service (1, 3). Third, “edge providers,” such as Google or Amazon, provide content, services, and applications over the Internet. And fourth, end users consume edge provider’s content, services, and applications.

As a simplified example of the Internet’s architecture, when an edge provider such as YouTube transmits content to an end user, that content is broken down into packets of information that are carried by the edge provider’s local access provider to the backbone network. This network then transmits these packets to the end user’s local access provider, which transmits these packets to the end user. Proponents of network neutrality focus on the relationship between broadband providers and edge providers. They fear that broadband providers might prevent their end-user subscribers from accessing certain edge providers altogether, or might degrade the quality of end-user subscribers access to certain edge providers, either as a means of favoring their own competing content or services or to enable them to collect fees from certain edge providers (1,4). Providers of internet resources can restrict or disrupt access to web content and applications in a variety of ways, including discriminating against competitors’ services and limiting the diversity of content (7).

A provider could discriminate by making sure that preferred content or applications load faster and more efficiently while competing services are slow or spotty. This practice would effectively create a tiered Internet – a fast lane for those who pay, and a slow lane for everyone else. To limit content, a provider can enhance its own web content and services by featuring prominent menus, program guides, start screens, etc. while systematically excluding other provider’s content. In response to these potential disruptions of Internet service, Vinton Cerf, considered a “father of the Internet,” as well as Tim Berners-Lee, the creator of the Web, and many others have spoken in favor of net neutrality and against the idea that network providers could choose to discriminate and decide how fast data could be transmitted and at what quality (1,4). In contrast, opponents of net neutrality claim that broadband service providers have no plans to block content or degrade network performance. Critics of net neutrality also argue that data discrimination of some kinds, particularly to guarantee quality of service, is not problematic but is actually the best way to encourage greater competition among such providers, which is currently limited in many areas (8). In response, net neutrality proponents claim that telecom companies seek to remove competition, create artificial scarcity, and oblige subscribers to buy their otherwise uncompetitive services.

**Regulating Net Neutrality**

As originally developed, the Internet was mainly an unregulated medium whose openness led to numerous innovative applications and services. The 1996 Telecommunications Act marked the first time that the Internet was included in broadcasting and spectrum allotment. In enacting it, Congress did not want the Federal Communications Commission (FCC) to treat websites and other Internet services the same way it treated the local access networks that enable people to get online. Congress understood that the owners of the access networks have tremendous gatekeeper power and so required the FCC to treat these network owners as “common carriers,” meaning they could not block or discriminate against the content that flows across their network to/from a person’s computer (1,2,3).

In response to the Telecommunications Act, ISPs began to explore alternative methods of network management that would permit them to discriminate the quality and speed of access to online content as they saw fit (2). This spurred the FCC to enshrine “net neutrality” in regulatory policy as a means of preserving the Internet’s open, non-discriminatory characteristics. However, in a reversal of policy, the FCC later decided to classify broadband Internet access service as an “information service,” meaning that broadband Internet would be subject to less regulation and that Internet access would be viewed no differently from a website or an online service. These regulatory decisions removed the FCC’s ability to prohibit ISPs from blocking or discriminating against online content.

“The 1996 Telecommunications Act marked the first time that the Internet was included in broadcasting and spectrum allotment. In enacting it, Congress did not want the Federal Communications Commission (FCC) to treat websites and other Internet services the same way it treated the local access networks that enable people to get online.”
Since changes in FCC policy, the neutrality of the Internet has been subject to fierce debate in the public arena and the subject of legal disputes. In 2014, an important case on network neutrality was decided by the U.S. Court of Appeals for the D.C. In Verizon v. FCC, the court struck down the FCCs “Open Internet Order” (promulgated in 2010) which was a set of regulations that advocated the policy of net neutrality (9). The court in Verizon stated that the FCC lacked authority to regulate net neutrality because of “the Commission’s still-binding decision to classify broadband providers as information services” (2,9). Some commentators argued that as a result of the court’s decision, Internet service providers can now block any website or application they want (see generally 9,10). However, while the court did recognize the statutory authority of the FCC to regulate broadband markets; it left open the door for the FCC to promulgate new rules to encourage net neutrality.

Conclusion

How can society ensure that the Internet remains open, accessible, and innovative, while encouraging the deployment of new broadband network resources? Some commentators have argued that net neutrality must be legally mandated least we lose the benefit of the internet. Others have advocated for less regulation and the reliance on market forces to determine the direction and freedom of the Internet (8). Even amongst those who argue that net neutrality is beneficial for free speech and unimpeded access to the Internet, most concede that some restrictions to the Internet are necessary. For example, Internet providers should be able to block spam emails, as well as viruses that could harm their networks and their customers’ computers. The road forward is in the hands of the FCC, the agency with the power to regulate net neutrality policy. Under the Verizon ruling, the FCC can still protect net neutrality if it is their intent to do so. The only thing left to clarify are the appropriate policies.

CONTACT JESSICA BARFIELD AT JESSICA.K.BARFIELD.15@DARTMOUTH.EDU

References

At icebreaker activities, there is one specific piece of trivia that I like to use whenever I have to give an interesting fact about myself: My full name is actually James Howe VI. To this day, I have never met anyone else who was also ‘the sixth’ or higher. James Howe I had a very interesting life: he was a two-time Congressman representing Queens and director of multiple banks in New York, and personally knew Theodore Roosevelt.

Does this have any bearing on my life? Could someone sequence my genome and show me that he had a significant genetic influence on me? For much of human history, many people might have answered: “yes.” From antiquity up until the Renaissance, noble families had pedigrees to show descent from notable sources, such as gods, legendary heroes, conquerors, or Roman patrician families, which they used to glorify themselves and add legitimacy to their rule (1). British inheritance laws also allow titles to be inherited by male-line descendants eight generations removed over any other more distantly related or female claimants (2). The 1896 Supreme Court case Plessy v. Ferguson, which established Jim Crow laws, revolved around an octoroon, or someone who is 7/8 white, man who was considered African-American (3).

Modern genetics has challenged these sorts of assumptions regarding the importance of one’s ancestry. Each person has two sets of chromosomes, one inherited from each parent. Knowing this, it can be estimated that an individual shares roughly half of its genes with each parent. There are other factors at play, like random mutation, but they do not have a large effect on the overall genome. In 1922, Sewall Wright, a researcher at the USDA, formalized and extrapolated this relationship by developing a measure known as the coefficient of relationship, commonly known as \( r \) (unrelated to the correlation coefficient) (4). The value of \( r \) is equal to the fraction of the genome that is similar to the family member of interest. It is a very simple calculation, where \( n \) equals the degrees of kinship separating the two related individuals (4). For example, parents are one generation apart from offspring, so the coefficient of relationship is \( 2^1 = 0.5 \), matching the observed phenomenon. Grandparents and great grandparents have \( r \)-values of 0.25 and 0.125, respectively (4). To answer my own question in the introduction, the \( r \) for between James Howe I and myself is 0.03125 (3.13% similarity,) so we are probably not very genetically similar at all.

As generations become more distant in time, the genetic similarity between members shrinks very quickly. In fact, when sperm banks do background checks on future donors to ensure that they carry no inherited diseases, the most stringent require medical history goes back only three generations, suggesting that any more distant individual genetic contribution is considered to be negligible (5).
some point a person’s genetic contribution to humanity actually becomes zero. At a separation of approximately 32 generations, an individual and its direct descendant of will have an \( r \) of 2.328 \( \times 10^{-10} \). The human genome is approximately 3 billion base pairs long (6). The product of these two numbers is less than one, meaning that such a distant ancestor is responsible for less than one base pair in its descendant’s genome, which is essentially equal to no contribution at all (7). In the United States, the average generation length is 25.2 years. Assuming that stays constant for the next 32 generations, it will take approximately 800 years for an individual to lose all impact in the gene pool (8).

**Identity by Descent**

However, while those living today may not have any discernable impact on the genes of anyone living in the year 2814, an individual’s genetic footprint remains enormous. Just as quickly as individual genetic contribution drops, the overall number of descendants rises, assuming these descendants all reproduce. To establish genetic distance between two living individuals, researchers use a method known as Identity By Descent (IBD). IBD relies on crossing over, an event during gamete formation where paternal and maternal homologous chromosomes exchange segments, creating recombinant chromosomes and, in effect, new gene profiles. To estimate genetic distance, geneticists search for sequences shared by two individuals at certain loci, known as the IBD block, inherited from a common ancestor (9). As the length of these sequences decreases, genetic distance increases, as more crossing over events have occurred (9).

Using this method, researchers at UC Davis were able to estimate the genetic distances between populations in European nations. Unsurprisingly, they found that people within nations were more closely related to each other than to individuals from other countries (10). They also found that Europeans had more distant cousins in close nations than in their own (10). For example, a German is most likely more closely related to a randomly selected German than to a randomly selected Greek, but he likely has more distant cousins living in Poland than in Germany (10). Interestingly, the group found that the common ancestor of all Europeans is much more recent than previously thought. Based on trends in decreasing IBD block length, they found that every single European shares a common ancestor who lived approximately 1,000 years ago, meaning every person of even partial European ancestry is a descendant of any European living before that time, including many of the first noble houses, such as the Carolingians (Charlemagne’s dynasty) and the Habsburgs (10,11). However, genealogists cannot conclusively determine deeper lineage at a more distant point in time. After the fall of the Roman Empire, the pedigrees of almost all notable Roman families ceased to be updated, and the Germanic tribes that took power in most of the former Western Roman Empire did not keep detailed records of ancestry (12).

**Royal Blood**

The establishment of European descent from antiquity, linking the earliest medieval nobles to Roman patrician families, is an important focus of genealogy today. By establishing a clear, verified line of descent between the two time periods, genealogists would be able to establish much longer pedigrees by combining records from the two time periods. The falsification of noble pedigrees in an attempt to enhance prestige has made this task significantly more difficult. Only a few individuals have been investigated thus far, and no families have been conclusively linked to anyone living before 800 CE (12). Europe is not unique in this regard. Most other parts of the world have similar difficulties establishing long lines of descent. Few genealogists have even attempted to create an African or Native American descent from antiquity, largely because of the destruction, loss, or lack of genealogy records in these regions. There are a select few families in the world that have produced a verified record of their descent from antiquity, with the most prominent three in Asia. The Yamato Dynasty, the Imperial House of Japan, can trace its roots back to the first historical Emperor of Japan, who reigned circa 300 CE (13). There is some debate over whether he was real or legend, however. While the Yamato Dynasty claims a fairly detailed pedigree, the longest verified pedigree belongs to the male-line descendants of Confucius. It stretches 79 generations to Confucius (551-479 BCE) and even continues further past the first king of the Shang Dynasty (b. circa 1700 BCE) (14, 15).

Another intriguing group includes the descendants of Muhammad. Detailed records of the lineage of some of his descendants exist, and numerous clans and royal houses in the Arab World claim descent from Muhammad (16). However, the claimants vastly outnumber the actual descendant due to the special place afforded to Muhammad’s descendants in society, where they are given the title of ‘Sayyid.’ This title affords them a number of privileges in Muslim society, such as prestige, specially colored turbans, and exemptions from various taxes (16). At one point, almost all of the Ottoman aristocracy claimed to be Sayyids to capitalize
on the myriad benefits that accompanied this lineage (16). Without genetic evidence, which members of most descendants from Muhammad are hesitant to provide, it is impossible to verify the ancestry of the millions of claimed “Sayyids” in the Arab world. Not all descendants of Muhammad are Arab, or even Muslim. In fact, there were three dozen verified descendants of Muhammad living in Spain who interacted with the common population, which, returning to the IBD studies from before, makes their common ancestors of all Europeans (17).

However, Muhammad’s descendants are not special in this regard. This phenomenon applies to almost every historical figure from the distant past. Most distantly historical figures have numerous descendants. For example, Genghis Khan, who almost every ruler in Asia traced descent from in order to gain legitimacy, had a known mutation on one of his Y-chromosomes that 8 percent of all men of Asian descent carry today (19). However, if women and female-line descendants had a similar marker, the number would be even higher. According to Joseph Chang, Professor of Statistics at Yale University, the most recent common ancestor of all of humanity, barring completely isolated populations, is thought to have existed around 3,000 years ago, also making every single human a descendant of any ancient figure living before 1000 BCE (assuming they had offspring whose lineage survived to the present day) (19). When two people are separated by more than four degrees of kinship horizontally, they are generally considered dissimilar enough not to be related. Ancestors are still considered a part of a person’s family, but four degrees of kinship away from your ancestors is the same degree of genetic similarity to four degrees of kinship away from cousins. It is truly odd that some people will take pride in and be given some social advantages based on the actions of their great-great-grandparent (who, based on genetic similarity, they do not resemble at all), but not their third cousin, who is just as closely related. When talking about a person, it can be said that his distant lineage has absolutely no effect on him on a personal level.

**References**

Introduction

Communications within networks of organisms is common among animals — humans, birds, and other social creatures — all constantly communicate a stream of information. For plants, however, methods of communication such as calls or body gestures are infeasible. Instead, this communication occurs primarily through detection and release of infochemicals (chemicals that carry chemical information) which must traverse the soil matrix to between roots or be emitted into the air as semi-volatile chemicals (1,2).

Although air-borne infochemicals are fairly well understood, recent research has focused on underground networks. Though this transfer of information may happen directly between plant roots, common mycorrhizal networks (CMNs) — networks of underground fungi — have been proven to greatly increase the ability of these chemicals to travel through the soil medium (1). It is difficult for infochemicals to simply diffuse through the soil to reach the intended destination due to the random nature of diffusion. However, fungal pathways made up of arbuscular mycorrhizae (a type of filamentous fungus) can act as superhighways of infochemical information (1).

Worldwide, most plants live in symbiosis with arbuscular mycorrhizal fungi (AMF) or the phylum Glomeromycota (3). CMNs facilitate community dynamics between plants, the mycorrhizae, and other underground organisms. Thus, plants may systemically signal to one another to defend against threats or otherwise react to environmental stimuli (1).

Communication through a Soil Medium

“Infochemical” is a broad term encompassing substances, including plant hormones and secondary metabolites (chemicals not directly involved in the growth and development of the plant) used in communication between plants, microbes, and soil animals below ground. Infochemicals mediate many types of communication between plants and their neighbors, including parasitic plants, plant growth-promoting bacteria, and other soil organisms (1).

In order for the message sent by an infochemical to be “received,” the “sender” must release it in sufficient quantities for it to successfully traverse the soil matrix and trigger a response. This presents a problem, as the initial concentrations of infochemicals released decline rapidly due to degradation (1).

Infochemicals are generally not highly chemically stable and can be degraded by absorption or adsorption to organic matter, biotic and abiotic decomposition, and the formation of metal complexes (1). If an infochemical is released and is not detected by the intended “receiving” organism, then it was not released in a sufficient level to become biologically effective (1).

The network enhanced bioactive zone (NEBaZ) model of CMN proposes that CMN facilitate rapid infochemical transport through the soil and also protect the chemicals from degradation, sorption, and complex formation (1). Because the chemicals are transported along the fungal hyphae (“roots”) along water...
potential gradients, they can be moved more rapidly and more safely than if they were left to diffuse through soil on their own (1).

**Common Mycorrhizal Networks**

Many filamentous microbes inhabit the soil, including the mycorrhizal fungi that are so useful to plant communication (1). Mycorrhizal fungi (fungi symbiotic with a plant’s roots), certain pathogens, and endophytic fungi – a fungus that lives within the plant, in this case in roots – all create networks between plants. However, only mycorrhizal fungi, particularly arbuscular mycorrhizal fungi (AMF), have been evidenced to transport infochemicals (1).

Mycorrhizal fungi evolved this function to better their own chances of thriving. As obligate symbiotes with the plants within their networks, the plants and the fungi mutually rely on one another for survival (1). Therefore it is thus unsurprising that CMNs have low specificity for hosts and are connected to multiple host plants. If one dies, the AMF does not necessarily perish along with it and can derive carbon from its other hosts (1,4).

Though the plants collect carbon aboveground via photosynthesis, the mycorrhizal hyphae are capable of taking up nutrients containing phosphorus and nitrogen and delivering them to the host plants (1,4). The hyphae may even assist in transferring the phosphorus and nitrogen to plants that need them the most (4). However, water flows in both directions; the host plant requires more water during the day (due to transpiration,) and the soil containing the mycorrhizal hyphae requires more water at night (1).

In addition to the water that travels within the hyphae of the AMF itself, a layer of water flows along the outside of the hyphae that actually may transport a greater volume of water. This phenomenon results from water potential gradients along the hyphal exteriors (1). Both represent viable methods of infochemical transport. In total, five possible routes of infochemical transport via CMNs exist.

First, the chemicals may be transported cytoplasmically across fungal cell membranes. However, this requires membrane transporters that allow the passage of organic infochemicals, and no evidence has yet been found of these sorts of hyphal membrane transporters (1). Second, infochemicals may diffuse through the fungal cell wall. However, this is only an option for nonpolar infochemicals given the hydrophobic nature of the cell wall (1).

Third, infochemicals may dissolve in the layer of water on the surface of the hyphae. This pathway is ideal for water-soluble infochemicals (1). Fourth, mycorrhizal hyphae twine together to form cords upon exiting plant roots. Within these cords, hydrophilic infochemicals may be carried while dissolved within water (1). Finally, the infochemicals may simple diffuse through the soil, though the hyphal activity within the soil generally facilitates this simple diffusion through soil aggregation, conductivity, and modification of the microbial community (1).

**Significance to Plant Science**

In addition to providing a conduit for nitrogen- and phosphorus-based nutrients, CMNs aid in plant defense and chemical production by allowing for the transfer of information amongst small communities of plants. They allow plants to transmit information to avoid disease and other threats. Thus, the AMF protect their hosts, and the plants protect one another (i.e. a diseased plant in the vicinity will likely impact its neighbors) (1).

For example, potted tomato plants infected with *Alternaria solani*, a fungal pathogen known to devastate tomato plants showed evidence of communication between intact CMNs with healthy neighbors. Uninfected plants in the vicinity showed evidence of increasing production of several defense-related proteins and creating greater disease resistance as a result. For trials with no active CMN between plants, there was little to no increase in disease resistance (5). Normally, tomato plants are spaced to minimize competition of plants for sunlight and nutrients. However, these results suggest that plants may actually benefit to growing in closer proximity. The benefits of plant defense through CMN may be balanced with the costs of plant competition to increase agricultural productivity in the long run (1).

In addition, Simard in 2009 suggests that CMN may serve a foundational role in allowing for Douglas-fir trees to act as a “hub” for the forest system. These trees even out resource availability with the networks and create

---

**Figure 2:** Found ubiquitously, arbuscular mycorrhizal fungi are able to transport infochemicals, nutrients, and water throughout its network of hyphae.
favorable conditions for the establishment of new trees. For populations of trees such as the Douglas-fir that require conservation and restoration, CMNs serve a key role in maintaining ecological favorability (6). In fact, Buscardo notes that CMNs help to repopulate forests after fires by connecting seedlings to access additional nutrient pools and by promoting water sharing (7).

In 2012, Simard explains how mycorrhizal networks enhance the rate of growth of a shrub throughout the arctic tundra. Because of the warming climate, the spread of the shrub *Betula nana* allows for the transfer of greater quantities of carbon to existing CMNs. As the CMN grows, greater quantities of phosphorus and nitrogen can be fed back to the shrub. A positive feedback loop occurs and amplifies the ecosystem change (8).

All of these instances suggest that CMNs can be represented through complex adaptive system theory. In complex adaptive systems, ecosystems are modeled as dynamic networks of parts interacting to create interactions that feed into self-organization and emergent properties (8).

CMNs provide direct avenues through which these interactions occur. They are central to the organization of most terrestrial ecosystems due to their foundational role in plant establishment. CMNs, interestingly, function from a bottom-up, localized system of interactions that iteratively links together. The behaviors of each part of the network influences the entire CMN (8).

**Conclusion**

Common mycorrhizal networks are still not fully understood, though with increased molecular and genetic tools, understanding of CMNs has increased greatly in the past five years (8). In the future, three main areas require further research: development of new analytical tools to detect and quantify infochemical concentrations, experimental designs to explain the underlying mechanisms of the NEBaZ model, and evaluation of CMNs ecological relevance (1).

Plant ecology has historically been viewed with an aboveground bias. Admittedly, plant communication may also occur aboveground via plant-emitted semi-volatiles that also release infochemicals (2). CMNs bring a necessary focus to underground ecology and the interconnected nature of the ecological community. There is strong evidence that plants may help other plants indirectly via CMNs (4).

Common mycorrhizal networks are fundamental agents of complex ecosystems, as they allow for dynamic feedback and cross-scale interactions. In addition, they play an enormous role in the organization of every ecosystem containing plants around the world. Whether in agricultural, conservationist, or purely scientific contexts, common mycorrhizal networks will remain in the spotlight in plant ecology in the coming years.

**References**

2. S. J. Himanen, J. D. Blande, J. K. Holopainen, Plant-emitted semi-volatiles shape the infochemical environment and herbivore resistance of heterospecific neighbors. Plant Signaling & Behavior 5, 1234-1236 (2010); published online EpUb10/01 (http://dx.doi.org/10.1089/jsb.2009.0899).

CONTACT YI (ANNIE) SUN AT YI.SUN.16@DARTMOUTH.EDU

---

**Figure 3:** Simple diffusion of infochemicals through the soil between different plants’ roots is often insufficient. Common mycorrhizal networks help to expand the bioactive zones of the infochemicals.
Introduction

Knitting together education, entertainment, and social networking, Massive Open Online Courses (MOOCs) can be described as the bridge between technology and education. MOOCs are more than online courses. They both transport the best education to remote corners of the world and help people advance their careers and expand their intellectual and personal networks.

Massive Open Online Courses

As the name suggests, MOOCs are online classes geared toward a large number of students with minimal professor involvement. Unlike traditional college courses, which carry credit, charge tuition, and limit enrollment, MOOCs are free, credit-less, and massive, with class sizes ranging from 100 to 100,000 students. Although there are currently no institutions that offer credit for their MOOCs, participants can buy or receive certifications confirming their understanding of the material (1).

MOOCs are also open to the public, meaning that anyone with an Internet connection can enroll. There is no registration cost and no admissions process — all one needs is a username and a password. In these courses, students watch short video lectures, usually about eight to twelve minutes long, which typically pause twice for a quiz — or in the case of computer programming, to allow viewers to code — in order to ensure that participants understand the material. Additionally, there may be homework, an online discussion board that is facilitated and monitored by teaching assistants, and a final exam. All assignments are submitted electronically and either graded by machines or other students afterwards, allowing professors to support courses with thousands (2).

Additionally, MOOCs are not limited to colleges and universities. A number of cultural organizations and philanthropies also sponsor MOOCs, offering short classes on topics related to the organization’s expertise (3).

History and Precursors

The first MOOCs originated from the open educational resources movement, which provided freely accessible, openly licensed documents and media for teaching, learning, and research purposes. At the time, the idea to provide free academic knowledge to the public and the premise of online courses was not entirely novel. For example, the Massachusetts Institute of Technology program, OpenCourseWare, had been giving people access to university lectures for almost fifteen years. Online models of teaching had also been available for a number of years at colleges and universities and provided credit or degrees for tuition-paying students (1).

Another fundamental precursor was Khan Academy, the non-profit, educational website founded by Salman Khan in 2006. Although Khan’s business originated with short, math tutorial videos for his younger cousins, it has since attracted millions of students worldwide. Today, it provides thousands of video lectures and exercises on a variety of subjects including mathematics, economics, and history, via YouTube (3).

Following the birth of Khan Academy, Apple launched iTunes U in 2007, which offers
Many universities and colleges joined after its inception and created courses designed for the online format by posting podcasts, video lectures or textbooks on the site for free download (3). During that time (2).

Coursera, a for-profit company founded by Stanford University computer science professors Andrew Ng and Daphne Koller, first partnered with Stanford University, Princeton University, the University of Michigan, and the University of Pennsylvania in April 2012. Since then, Coursera has signed contracts with a number of elite universities and colleges around the world. One of the main benefits of MOOCs stems from the flexibility of the courses. Although each platform differs in its time constraints, each allows users to proceed at their own pace. For example, in Udacity, viewers can start a class whenever they wish and complete a task or exam at their own speed, reducing the size. In contrast, Coursera lists a start and end date for their classes and due dates for exams and assignments. At the same time, students are able to watch lectures at any time and can pause, rewind, and comment on the videos (3).

Benefits

The year 2012 was deemed “The Year of the MOOC,” as many of the major platforms such as Coursera, edX, and Udacity were launched during that time (2). These agreements stipulate that in exchange for using Coursera’s platform to offer free courses, the schools will receive a percentage of the revenue (1). Described as a hub for learning and networking, Coursera currently boasts 633 courses, ranging from classes in computer science and medicine to philosophy and music, and 108 elite partners including the University of Toronto, Peking University, and Yale University (5). Although Coursera can offer suggestions, the institutions make all of the pedagogical decisions, mainly adapting their online offerings from their existing courses (2).

edX, which provides a similar platform to Coursera, was jointly launched in May 2012 by the Massachusetts Institute of Technology and Harvard University. Unlike Coursera, edX is a non-profit organization supported by $30 million contributed by each institution. Currently, edX offers over 150 courses in the humanities, math, computer science, and physics and employs over 400 faculty and staff members who are responsible for teaching courses and discussing topics online. Today, edX has more than 2 million users that have earned over 100,000 certifications (6).

A third popular platform is the for-profit company Udacity, which was founded by Stanford University computer science professor Sebastian Thurn. Instead of working with elite institutions, Udacity pairs with individual professors. Its courses, which are designed and produced in-house or with companies like Google or Microsoft, focus exclusively on computer science and related fields, with a focus on applied learning and hands-on skills (2).

edX, which provides a similar platform to Coursera, was jointly launched in May 2012 by the Massachusetts Institute of Technology and Harvard University. Unlike Coursera, edX is a non-profit organization supported by $30 million contributed by each institution. Currently, edX offers over 150 courses in the humanities, math, computer science, and physics and employs over 400 faculty and staff members who are responsible for teaching courses and discussing topics online. Today, edX has more than 2 million users that have earned over 100,000 certifications (6).
evaluation of the pedagogical foundations of MOOCs, David Glance and his research team found that benefits could be seen in almost all characteristics. For example, the online quizzes and assessments enhance retrieval learning, the short videos and quizzes help mastery learning, the short videos allow improved attention and focus, and the online mode of delivery reinforces the efficacy of online learning (10).

### Challenges and Criticism

Despite their many benefits, MOOCs also face criticisms. One of the most noticeable characteristics of MOOCs is their low completion rate. Although thousands of students may sign up for a course, less than 10 percent will complete it, with a steep participation drop starting in the first week. For example, in the course “Bioelectricity” at Duke University in the fall of 2012, 12,725 students originally enrolled in the course. Of those participants, only 7,761 watched a video, 3,658 attempted a quiz, 345 attempted the final exam, and, ultimately, 313 passed and received a certificate (11). From Coursera shows similar results, with a completion rate of seven to nine percent (12).

One online survey listed factors such as timing, lack of prerequisite knowledge, lecture fatigue, and poor course design as popular reasons for dropping a course (13). Although MOOCs show a great deal of promise in expanding educational opportunities, there are still many kinks that need to be addressed. For example, course placement is complex, as it is difficult to discern which course is right for an individual. While MOOCs are beneficial if one is simply looking for knowledge, the situation grows more complicated if participants are seeking credit. As of now, it is difficult to understand how these certifications will fit into the structure of higher education (2).

An additional challenge involves making the massive nature of MOOCs more intimate. As in all courses, one of the most important components is instructor connection and student interaction. Although educators and researchers are trying to strengthen these relationships, they have proven difficult to resolve. For example, on Coursera, users have the opportunity to build a profile, where they can upload a photo and utilize tools to plan social events or study groups with fellow students in 1,400 cities worldwide (5). The success of these meet-ups depends in large part on the initiative taken on by the members. Another concern is the fairness of the grading system. For the assignments that cannot be scored by an automated computer, Coursera utilizes peer grading, especially in courses involving writing and analysis. After submitting an assignment, five people will read and grade it and, in return, the writer will grade five other assignments. This system of grading can have potentially unfair, if not disastrous, consequences (2).

### The Future of MOOCs

Ultimately, MOOCs raise fundamental questions about the future of teaching, the value of a degree, and technology’s long-term effect on universities. Although the system is still far from perfect, by using different platforms like forums, blogs, and social networks, these courses allow students to create a personal, independent learning environment and gain exposure to interconnected knowledge.

**References**

12. D. Koller, Interview with Knowledge @ Wharton, 7 November 2012.
Network Science: The Future of Medicine?

At the dawn of the new millennium in January 2000, Stephen Hawking was asked about his outlook on the future of science. "I think the next century will be the century of complexity," the renowned Cambridge physicist predicted (1).

Indeed, a rapidly growing population and breakthrough technological advances in the last decade have created a world more complex than ever. From the convoluted web of international trade to the global network of social connections made possible by the Internet, complex systems have arisen in many aspects of every day life. The variables within a complex system interact in interdependent and often unpredictable ways. Because of this, the traditional reductionist paradigm, in which scientists try to understand and describe a system by studying its component parts, is not adequate (1). In order to gather a full understanding of the functional interdependencies of a complex system, many scientists have adopted a more complete, holistic approach in their research.

Part of the holistic approach has involved building networks as a tool to visualize the basic characteristics and principles of physical, biological, and social systems. In the last ten years, network science has emerged as a promising new field. As network theory continues to be developed and refined, the biomedical community has already recognized the many potential biological and clinical applications of network-based approaches to human disease. Using network science to study human disease, an approach termed "network medicine," has led to a greater understanding of complex diseases, as well as the exciting potential of novel therapeutic strategies to fight these diseases.

What is a Network?

Mapping a set of system components called "nodes" generates a network, and "links" or "edges" represent an interaction of some kind that connect these nodes. Many different types of biological networks have been constructed and analyzed, including protein-protein networks, whose nodes are proteins that are linked to each other by physical interactions; metabolic networks, whose nodes are metabolites that are linked if they participate in the same biochemical reactions; and human disease networks, whose nodes represent human genetic disorders that are linked if they share a mutated gene (4).

The earliest models of complex systems were random networks. In these models, a fixed group of random nodes are selected and randomly connected to other nodes. According

Figure 1: Network models like these show the complexity of human disease by mapping the vast number of interactions happening at a molecular level. They will be critical in helping future scientists understand the mechanisms of complex diseases.

Image courtesy of Jingya Qiu
to the random model, a network forms when the system reaches a critical number of connections, which means its formation is a sudden occurrence, not a gradual one. Over the last few decades, mathematicians and physicists have discovered that the random model does not apply to networks in the real world. Real-world networks do not start with a fixed number of nodes and connect randomly. Rather, they start with a small nucleus and add nodes one at a time, preferring to connect to already highly connected nodes (5).

Consider one of the great human-constructed networks of the 21st century, the World Wide Web. Internet did not start out with a fixed number of web pages but with a small nucleus of web pages and one by one increased the number of web pages until it became the behemoth it is today. Furthermore, the probability that two web pages are connected is not uniformly random. A newly created webpage will show preferential connectivity towards web pages that are already highly connected because they are more likely to be known (5). These two properties of gradual growth and preferential connectivity led to the conclusion that the World Wide Web can most accurately be represented by a "scale-free" network model, not a random network model.

Incredibly, almost all real-world networks, both human-constructed and social, are united by the same underlying set of organizing principles. Technological systems such as electronic circuits, social systems such as email, the web, and natural systems such as protein interactions, brain function, and earthquakes can all be characterized by scale-free networks. The ability to apply general network theory to biological networks has allowed scientists to build more accurate and representative biological models.

The degree of a node is the number of links that are connected to it, and the degree distribution of a network shows the probability that each node will have a certain degree (6). While random networks’ degree distribution follows a normal bell-curve distribution, scale-free networks’ degree distribution is a right-skewed plot with a long tail. In other words, most of the nodes in a random network have an average number of links while the majority of nodes in a scale-free network are lowly linked with a few highly linked nodes called “hubs” (6).

Because scale-free networks are around hubs, they are very robust to random failure. In fact, if 99 percent of nodes are removed randomly from a scale-free network, the remaining one percent of nodes can still remain intact. The downside to scale-free networks, however, is they are very fragile to targeted attacks. If hubs are removed specifically, the network breaks apart very quickly. In biological networks, hubs correlate very strongly with essential genes that code for proteins without which humans cannot be born (7). Most disease genes are represented in lowly connected nodes on the periphery of networks. Evolutionarily, this makes sense because organisms generally cannot live long enough to pass on a mutation that significantly affects the function of an essential protein (7).

**Network Medicine**

Traditionally, the most common method of studying the genetics of a disease has been to test if a single genetic variant correlates with an increased risk of the disease. This method of studying genes in isolation is a dramatic oversimplification. It ignores the existence of continuous interactions between proteins, gene products, and different signals in the metabolic pathways. Human disease is rarely the result of a mutation in a single genetic variant, but rather, the net result of many macromolecular complexes that interact through an interconnected network (2). At best, the simple genotype-phenotype approach is only able to explain about twenty percent of disease risk due to genetic variation (3). Many network scientists believe that in order to account for the other eighty percent of disease risk, it is necessary to understand genes in the context of intricate biological interactions on the molecular level.

Disease genes are not randomly positioned in the biological network. Network scientists hypothesize that genes or molecules involved in the same biochemical process or disease have a much higher tendency to interact with one another. In other words, biological networks are modular. This suggests that if a few disease components are identified, other components of the same disease are likely to be found in the same neighborhood, often referred to as a “module” (4). Disease genes can be studied with three different types of modules. Topological modules show nodes that have a higher tendency to link to nodes within the same local neighborhood. A functional module shows aggregations of nodes that are involved in similar functions. A disease module shows a group of network components that contributes to a disease phenotype (4). Cellular components that belong in the same topological, functional, or disease module have a higher likelihood of being involved in the same disease. Because of this property, modules are used in conjunction with bioinformatics techniques to find previously unknown genes associated with diseases and predict disease risk (8). Furthermore, modules also have the power to narrow and simplify the enormous complexity of the overall network. This gives biologists searching for disease mechanisms and
drug therapies a greatly reduced area to focus on, easing their studies significantly.

In 2007, Goh and a team of scientists at Notre Dame used already existing gene-disease associations to build the human disease network (HDN). Diseases were linked if they shared one or more genes, indicating a common genetic origin. As expected, this network followed the scale-free model and showed clusters of nodes forming modules. An analysis of the HDN revealed that a patient is twice as likely to develop a particular disease if that disease shares a gene with the patient’s primary disease (7). This suggests that shared genetic origin greatly amplifies the chances of comorbidity. The HDN also revealed the unique power networks have in revealing underlying connections between diseases that ordinarily would not be seen (6). Two diseases that might seem completely unrelated currently may prove to share a common gene. If this is the case, they also share a common genetic origin and, most likely, a biochemical pathway.

**Network Pharmacology**

The current paradigm of drug discovery is gene and protein-centric (9). Scientists identify specific proteins involved in a disease and design drugs to specifically target one protein, hoping it will lead to the eventual curing of the disease. There are many disadvantages to this approach, but perhaps the most significant one is unwanted side effects. Drugs that are supposed to bind to a single target may correct some functional aspect of the disease, but they could also end up having multiple interactions with other molecules within the disease module, which would change the efficacy of the drug. Recent drug development efforts have started
using network biology approaches to identify protein-protein interactions and other points for therapeutic intervention. As a result, the paradigm of drug discovery has shifted towards a network and pathway-centric approach, taking into account the biological context of drug targets.

The network approach to drug development has resulted in an increased effort to identify protein-protein interactions suitable for drug therapy. Targeting interactions rather than more traditional targets such as enzyme active sites helps to avoid unintended side effects (9). Biological networks have also led to the belief that most disease phenotypes cannot be cured with a single drug. Thus, increased efforts have also gone into developing drug therapies with multiple targets. These combinatorial therapies have already been used in treating AIDS, cancer, and depression (4). In future years, the ultimate goal will be to develop methods of identifying the optimal drug combinations from networks.

The Future

Though network medicine shows promising potential, the approach is still greatly limited by the incompleteness of current biological networks. Current protein interaction networks only show about twenty percent of the existing links (10). Even with the existing incomplete network, there are many applicable uses. Functional modules can be constructed, protein-protein interactions can be identified, and diseases and comorbidities can be predicted. Yet, without eighty percent of the links, many patterns are invisible and a lot of valuable information is lost.

In a lecture given in February 2013, Albert-Laszlo Barabasi, current Professor of Network Science and Director of the Center for Complex Network Research at Northeastern University, compared the job of an auto mechanic to that of a physician. An auto mechanic has the capability of fixing virtually any car problem because he has a list of each car part, as well as a blueprint of how the parts wire together. The Human Genome Project gave doctors “the parts” to human disease. But a doctor with just a list of genes in isolation is a mechanic without a blueprint. The critical next step is to construct a blueprint for these parts and to visualize how genes interact with each other to produce each phenotype (10).

Ten years ago, network science resided almost exclusively in physics departments, garnering very little attention from other fields. In the last decade, as scientists realize the ubiquitous nature of networks and the practicality of their applications, network science has grown rapidly in popularity. It has transcended physics to the realms of computer science, biology, neuroscience, and sociology. Medical doctors have also realized the great importance of network science. Harvard Medical School just recently established a division of network medicine and Geisel School of Medicine at Dartmouth has plans for a visualization lab for network science in their new research building.

Perhaps most importantly, network medicine has transformed the way scientists and doctors approach human disease. With the knowledge amassed from networks, it is clear that the future lies in embracing the complexity and interconnected nature of disease rather than ignoring it. For a long time, doctors have thought of human disease as genetic mutations in isolation. Network medicine calls for doctors to let go of this reductionist paradigm and think about the disease within a larger system of molecular biology.

CONTACT JINGYA QIU AT JINGYA.QIU.16@DARTMOUTH.EDU

References

3. J. Moore, “Bioinformatics: 25 Years of Integrating the Biological Sciences” (31 March 2014). Lecture Delivered at the 26th Presidential Faculty Lecture, Dartmouth College, Hanover, NH, USA.
Introduction

While anorexia nervosa is often viewed solely as an eating disorder, new research suggests that a neural framework similar to that seen in individuals with autism underlies the surface behaviors of the disease. Anorexia nervosa (AN) is characterized by the persistent refusal to eat, a behavior fueled by a pathological fear of gaining weight and a skewed perception of one’s body image. Individuals with anorexia place undue importance on the need to be thin and often fail to recognize the seriousness of their condition (1). Anorexia is nearly four times more common in females than in males, with an average prevalence of 0.3 percent nationwide (2). Autism spectrum disorder (ASD), on the other hand, is a neurodevelopmental disorder characterized by restricted and repetitive behaviors and deficits in communication and interpersonal interactions. Autistic individuals often focus internally, have difficulty understanding the emotional states of others, and demonstrate a need for sameness and routines. Autism manifests itself in a variety of ways and is best described as a group of conditions presenting along a continuum (1).

Although anorexia and autism seem to present quite differently, they often exist simultaneously. According to a study from King’s College London, as many as 25.8 percent of females suffering from anorexia received a score that warranted further clinical evaluation on a diagnostic test for autism (2). While the proposal that anorexia should be classified on the autism spectrum was scoffed at 30 years ago, researchers recently discovered many social and cognitive similarities between the two disorders, although no neural correlates have yet been confirmed (2).

Overrepresentation of ASD in AN Populations

A groundbreaking study by Tchanturka et al. (2013) was the first of its magnitude to investigate the prevalence of autistic traits in AN females. Researchers administered the Autism Spectrum Questionnaire (AQ-10) to 66 female AN patients and 66 control subjects. The AQ-10 is a ten-question self-report survey adopted from the AQ-50, which is commonly used in clinical settings. Responses indicating autistic traits received a score of one, while all other responses received a score of zero. The scores from the ten questions were totaled, with a score of 6 out of 10 serving as the clinical cutoff for a potential ASD diagnosis. While only 1.5 percent of control subjects scored above 6, 25.8 percent of AN subjects received a score equal to or greater than the clinical cutoff. In addition, the mean score for AN subjects was 4.53, while the mean score for the control group was only 1.85. AN subjects also received higher average scores on each of the ten individual questions (3). This study shows that autistic traits are more prevalent and more severe in anorexic females than in otherwise healthy individuals.

ASD and Restrictive Eating Habits

Autistic individuals often seem to engage
in abnormal eating behaviors similar to those of individuals with anorexia. Caregivers of autistic children often report unusual eating behaviors including repetitive consumption of specific food types and/or brands and heightened sensitivity to the sensory qualities of food (3). As a result of these restrictive eating habits, children with autism eat fewer types of food within each food group and are often of lower than average body weight (4). Some pediatricians consider poor eating habits alone to be a reliable warning sign for autism (3).

**Similar Social Deficits**

Impaired social functioning is one of the primary diagnostic criteria for autism. While this is not true of anorexia, new research has revealed that AN individuals also have significant deficits in social processing. Furthermore, these deficits, which mainly surround emotional intelligence, seem to be quite similar to those seen in individuals with autism. Both clinical groups have difficulty recognizing emotional stimuli, expressing and describing emotions, experiencing empathy, and understanding the mental states of others. Deficits in emotional intelligence combined with social anhedonia (the inability to experience pleasure) may explain why AN and ASD individuals have difficulty maintaining interpersonal relationships (5).

The ‘reading the mind in the eyes’ task has been carried out with both ASD and AN patients to determine how well individuals in each clinical group can attribute mental states to others. In this well-known theory of mind test, participants are shown photographs of other people’s eye region and required to decide between given ‘mental state terms’ describing what the pictured person is likely to be thinking or feeling. Participants are awarded one point for each correctly chosen mental state term, so the maximum score is equal to the number of photographs shown (6).

A study conducted by Baron-Cohen et al. (1997) used the eyes task to compare the theory of mind abilities of healthy controls with two clinical groups: those with high-functioning autism or Asperger’s and those with Tourette’s syndrome. Performance on the task differed significantly between participants with ASD and control subjects, but not between subjects in the control group and subjects in the clinical group with Tourette’s syndrome. While the mean score on the task for healthy controls was 20.3, it was only 16.3 for ASD subjects. Also, none of the ASD participants received the maximum score of 25 (6). The results from this study show that autistic individuals are significantly less able to infer a person’s mental state based on facial expression than are other test subjects, suggesting compromised emotional intelligence.

In 2009, Russell, Schmidt, Doherty, Young, and Tchanturia were the first to utilize the eyes task to compare the performance of AN individuals with healthy controls. In this particular study, the total scores were reported as percentages of the maximum score of 36. While the mean score for healthy controls was 85.4 percent, the mean score for the AN participants was only 77.5 percent. Similar to individuals with autism, individuals with anorexia have difficulty using emotional information to guide thinking. Without the ability to appropriately perceive, interpret, and respond to the emotions of others, interpersonal relationships are understandably compromised.

**Similar Cognitive Deficits**

Deficits in cognitive processing are at the heart of ASD. ASD individuals experience decreased levels of cognitive flexibility and impaired global processing, abilities measured experimentally by set-shifting tasks and embedded figures tests respectively. While poor performance on such tasks has been associated with autism spectrum disorders for quite some time, new research has revealed that individuals with AN perform similarly on such tasks.

**Set-Shifting**

Set-shifting, the ability to modify ongoing behavior in response to changing goals, is a common measure of cognitive flexibility. The “cat bat task” tests subjects’ set-shifting abilities by asking them to fill in the missing letter when ‘_at’ appears in a story. In the first half of the story, the context requires that subjects add the letter ‘c’ to make the word ‘cat,’ while in the second half of the story the subjects must adjust their cognitive set because ‘bat’ becomes
the appropriate choice. While healthy women took 22.2 seconds on average to adjust to the shift, AN women performed similarly to ASD individuals taking an average of 30.7 seconds to adjust (7). AN and ASD individuals alike experience difficulty thinking about multiple concepts simultaneously. Decreased cognitive flexibility prevents the learner from adapting to changes in the environment, a skill required in the majority of daily activities.

Local-Processing

Designed to measure local processing abilities, embedded figure tasks (EFTs) ask participants to identify which of two complex shapes contains the given target shape. AN and ASD individuals are consistently more accurate in their decisions than are healthy controls, reflecting an unusual balance between local and global processing abilities (8). A study conducted by Ring et al. (1999) found that the demands of this task provoked very different patterns of brain activation in the clinical and control groups. While control subjects employed brain areas involved in visuospatial searching, such as the precuneus, AN and ASD subjects seemed to make use of an alternate cognitive strategy (8).

Conclusion

While anorexia and autism are currently separated in the DSM-V, the most recent edition of the Diagnostic and Statistical Manual of Mental Disorders, new research makes many researchers wonder if anorexia should be classified as a pseudo-autistic state or a female presentation of autism. As Dr. Bonnie Au yeung said, "Autism is diagnosed more often in males. This new research suggests that a proportion of females with autism may be being overlooked or misdiagnosed, because they present to clinics with anorexia" (9). This newfound link will hopefully provide insight into the etiology of AN.

The beginning stages of research suggest that AN, with its wide variety of deficits, may be better classified as a neurocognitive developmental disorder. However, some researchers do not believe that the prevalence of autistic traits in AN populations suggests similar neural underpinnings between AD and AN. Instead, they believe that the social and cognitive deficits seen in AN individuals are simply the result of starvation during critical developmental periods (3). Regardless, this new research has important treatment implications, as it provides clinicians with a novel framework to understand the minds of their AN patients. Perhaps clinicians should shift their focus from their patients’ restrictive eating habits to their unique neurocognitive styles. Considering that anorexia has the highest mortality rate of any mental disorder, a new approach to treatment is certainly in order.

CONTACT JULIA ISAACSON AT JULIA.E.ISAACSON.15@DARTMOUTH.EDU

References

Structural Holes, Brokers, and Fraud in a Clean Development Mechanism (CDM) Market

MATTHEW JIN

Abstract

Brokerage, which in this paper will be considered the cooperation of two entities on a project who otherwise would not work together, can be essential for the propagation of information in small world networks. For this reason, identifying the entities who facilitate brokerage, or brokers, is important for understanding a network’s structure. The network that will be considered in this paper is a Clean Development Mechanism (CDM) network, which is a network of entities that collaborate on emission-reducing projects, which earn Certified Emission Reduction (CER) credits, which can in turn be traded in a cap-and-trade scheme. Because the CDM network is a small world network—a network that is largely characterized by several clusters, with relatively few entities linking these clusters together—one may suspect that there are brokers in this network whose varied project experience make them important hubs of information flow—they would bring together entities in unlike clusters who would not typically work together and who would possess relatively specialized project knowledge. However, due to the reported presence of fraud within the CDM market, one may also suspect that those who appear to be brokers may in fact be those who are committing fraud. Therefore, the focus of this study is to identify the brokers in this market, to understand the extent to which these brokers are committing fraud, and to understand how both of these concepts relate to the market structure. After analyzing data from every CDM project up through February 2012, ten brokers were identified, all of whom were validators—entities which review and validate project proposals—or verifiers—entities which verify whether or not projects that have been implemented are creating the planned emission reductions. Many of these brokers were suspended for fraudulent activity by the Executive Board (EB) of the UN, and those who were not suspended received low performance ratings according to reports by the WWF. Moreover, evidence from the model shows that these brokers collaborate with each other to commit fraud. Overall the market appears to be driven in an oligopolistic manner by a few designated operational entities (DOEs), which is another term for an entity that is a validator or a verifier, which work with each other in committing fraud. While the model suggests that there is significant room for brokerage in the network, this brokerage is likely not occurring.

Introduction

The CDM is a flexibility mechanism introduced by the Kyoto Protocol in 2007. It allows industrialized countries to reach their emissions targets more cheaply by investing in sustainable development projects in other countries. There are typically three, and possibly four entities that work together in a single project: a host country, in which the development project takes place, a buyer country, which purchases the Certified Emission Reduction (CER) credits which the project produces, a validator, which independently evaluates the CDM project against CDM regulations, and a verifier, which periodically evaluates the project to certify that it is being implemented as planned and that it is properly monitoring its emissions. If helpful or necessary, a consultant may be brought in to facilitate the planning of the project.

Analysis of the CDM market suggests that it is structured like a small-world network. In a graphical representation of the network, a node is a point in the graph and an edge denotes a relationship between two nodes in the graph. In the context of the CDM market model, a node represents an entity of the CDM network and an edge denotes mutual participation in a CDM project. The nodes can be organized into ten “classes,” or groups whose members tend to collaborate with each other more than they would with members from other groups. This network has a relatively high modularity score of 0.366, implying that the network has dense connections within groups but relatively sparse connections between groups (1). Because of this, one is led to suspect that there are structural holes and therefore a potential for brokerage. In the case of the CDM market, brokers are entities who possess an unusually high amount of edges with entities outside of their own group. They could therefore bring wide and varied experience to the project and form helpful connections between entities that would otherwise not typically collaborate together.

The central questions of this study were to determine the brokers of this CDM network, to evaluate whether “brokerage” is at all indicative of fraud, and to evaluate how fraud and brokerage relate to the network structure. The primary tool in this investigation was MATLAB.

Upon analysis of the data, ten brokers were discovered among this network: DNV, SGS, TÜV-SÜD, RINA, LRQA, TUV-Nord, and BV Cert. For three different agencies—SGS, DNV, and TUV-Nord—both the validation and verification services were labeled as brokers. Two of them, SGS and DNV, were suspended for not adhering to inspection rules (2). TUV-Nord avoided suspension, yet earned a grade of D in a report by WWF and the Öko Institute.
for Applied Ecology that assessed to what extent the validators and verifiers fulfilled the requirements of the CDM Executive Board. In that report, SGS earned a grade of E and DNV received a grade of F—A was the best possible grade, while F was the worst (3). Nonetheless, of the over nine-hundred projects evaluated in the report, a D, which was given only once, was the highest grade awarded (3). According to the report’s rating scheme, a D meant that 50% of the projects proposed by the company were automatically registered, 25% were registered after corrections, 20% were registered after a review and corrections, and 5% were outright rejected by the EB. An F meant that 20% were automatically registered, 40% were registered after corrections, 20% were registered after a review and corrections, and 20% were outright rejected. The low ratings of these important brokers of the network thus indicate that many of their projects were being rejected, reviewed, or requested for corrective action by the EB. Thus, one might conclude that although there is almost certainly room for true brokerage in the network, and there are brokers with highly varied project expertise, the structural holes of the network are probably not actually being bridged due to the lackluster performance of the brokers.

Data

This study used a MATLAB record of all CDM projects up until February 2012. It included the name of the participating entities, the time duration of their participation in the CDM, and their roles (buyer, verifier, validator, or consultant). It also showed which entities in the CDM had mutual interaction in a project, and the duration of that relationship.

This study also used an Excel file containing a variety of other data about the projects. Among various other pieces of information, it listed the name of each project, the host country, other countries involved, the scale of the project, and the type of project. The MATLAB and Excel files came from the Institute for Global Environmental Strategies.

This data is applicable to this project primarily because it shows which entities worked in projects together, as well as the duration of each interaction. From this information, one can divide the CDM market into groups depending on the level of interaction, and from this one can determine the level to which each entity interacts with groups outside of its own.

Methods

Two adjacency matrices, which are matrices that represent which nodes of a graph are adjacent, were constructed from the data. One adjacency matrix was unweighted and undirected; a “1” entry denoted a mutual participation in a project for two entities. The other adjacency matrix was undirected, with each entry being weighted by the time duration of the relationship between the two entities normalized by the total amount of time that the entity spent interacting with other entities.

Using Gephi, this data was then organized by modularity class (1). Then, for each entity a modified out-degree was calculated, which in this paper will be called an “outscore.” It denotes the number of connections that the node has outside of its own cluster. One would expect a broker to have a high value for this statistic, because a broker that bridges structural holes would work with a variety of other entities outside of its own cluster.

For each entry a network constraint (NC) score was also calculated (4). For each relationship, \( c_{ij} = (p_{ij} + \Sigma p_{i\ell} p_{\ell j})^2 \), where \( p_{ij} \) is the proportion of time that entity i has worked in the CDM with entity j. The NC score for the entity is thus \( \Sigma c_{ij} \). Qualitatively, it represents the degree to which it connects entities which are already connected to each other. In other words, it is how tightly clustered each entity’s network of contacts is. One would expect a broker to have a low network constraint value, because one would expect it by definition to be linking contacts that are not closely linked together.

One would expect a broker to have a low network constraint but a high outscore. From Figure 1, one can see that the majority of the network constraints are below fifty, while one can see from Figure 2 that few outscores are above one-hundred. This study thus considers a network constraint lower than fifty to be low and an outscore greater than one-hundred to be high; a node which possesses both of these statistics is considered a broker.

Results

This study identified ten brokers in the CDM market. Various statistics about these brokers are detailed in the table on the right. Note that % degree out denotes the percentage of the node’s total out-degree that is represented by the outscore.
Table 6.1: Statistics for ten brokers in the CDM market. Note that % degree out indicates the percentage of the node's total out-degree that is represented by the outscore.

**Table 6.1: Statistics for ten brokers in the CDM market.**

<table>
<thead>
<tr>
<th>Broker Name, Role</th>
<th>Outscore</th>
<th>NC</th>
<th>Degree</th>
<th>% Degree Out</th>
<th>Modularity Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNV, validator</td>
<td>138</td>
<td>19</td>
<td>231</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>DNV, verifier</td>
<td>687</td>
<td>8</td>
<td>998</td>
<td>69</td>
<td>9</td>
</tr>
<tr>
<td>SGS, validator</td>
<td>187</td>
<td>14</td>
<td>277</td>
<td>68</td>
<td>1</td>
</tr>
<tr>
<td>SGS, verifier</td>
<td>405</td>
<td>1</td>
<td>677</td>
<td>60</td>
<td>8</td>
</tr>
<tr>
<td>LRQA, verifier</td>
<td>103</td>
<td>13</td>
<td>183</td>
<td>56</td>
<td>9</td>
</tr>
<tr>
<td>TÜV-SÜD, validator</td>
<td>112</td>
<td>41</td>
<td>192</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td>RINA, verifier</td>
<td>178</td>
<td>2</td>
<td>242</td>
<td>74</td>
<td>7</td>
</tr>
<tr>
<td>BV Cert, verifier</td>
<td>268</td>
<td>0.15</td>
<td>448</td>
<td>60</td>
<td>9</td>
</tr>
<tr>
<td>TÜV-Nord, validator</td>
<td>109</td>
<td>1</td>
<td>174</td>
<td>63</td>
<td>9</td>
</tr>
<tr>
<td>TÜV-Nord, verifier</td>
<td>374</td>
<td>9</td>
<td>688</td>
<td>56</td>
<td>9</td>
</tr>
</tbody>
</table>

Discussion

Of the fifteen outscores that were greater than one hundred, ten appeared as brokers in this network. In general, a high outscore was correlated with a low NC score. There were a couple of notable exceptions: China, with a network constraint value of 304 and an outscore of 332, and India, with a network constraint of 556 and an outscore of 372. Note that in spite of the high raw outscore, these entities’ connections to modularity classes outside of their own made up a relatively small percentage of their total out-degree; for China it was 37%, while for India it was 35%. These statistics are an indication that even though one might expect a broker to have a large quantity of connections outside of its modularity class, those connections should make up a large part of the entity’s total out-degree in order for the entity to be considered a broker; the total out-degree for each of the brokers composed over half of the connections outside of the broker’s modularity class.

Even though one might expect China and India to be brokers due to the number of the projects that take place within their borders, their networks are highly constrained because they are geographic locations; local entities already work with each other out of geographical convenience, and hence the host country is not forming connections which did not previously exist to any significant extent.

There are two notable entities, TÜV-SÜD (verifier) and TÜV-Rhein (verifier), that have a relatively high proportion of connections outside of their own modularity class, yet still retain a high network constraint score. For TÜV-SÜD, the percent of total out degree outside of its modularity class is 82%, while for TÜV-Rhein it is 50%. TÜV-SÜD’s NC is 188, while for TÜV-Rhein it is 301. This suggests that although a significant proportion of the entity’s connections may be those outside of its modularity class, it is still possible for the network in which the entity operates to be constrained; there is only a correlative relationship between the two statistics.

All of the brokers were either verifiers or validators. This arises from the structure of the network (Figure 4). There are very few validators and verifiers relative to the amounts of entities occupying other roles. However, every project requires a validator and a verifier; thus, the market is virtually oligopolistic, and the few validators and verifiers in it must inevitably be connected to a wide variety of other entities.

This structure gives great opportunity for abuse. Because CER issuance is the principle source of income during a project’s

![Figure 3: India’s Network. Generated by Gephi. This picture is colored by modularity class. The central node is India; one can see that it is primarily connected with the yellow nodes within its own cluster.](image1)

![Figure 4: CDM Network Colored By Role. (Blue: consultant, Green: buyer, Yellow: host country, Red: verifier, White: validator)](image2)

Image courtesy of Scott Pauls, Dartmouth College, Math 76, Winter 2013. Lecture 5, Slide 16.
duration, validators and verifiers like DNV, TÜV-SÜD, and SGS are monetarily incentivized to verify and validate projects (2). Many of the approved projects were passed merely based upon the project developer’s word. As a result most projects are neither additional—they would have taken place anyway without help from the sales of the carbon credit—nor sustainable (4).

Ultimately, developed countries can therefore simply exploit the CDM as a way to avoid having to reduce their own emissions by claiming credits for fabricated emission reductions.

For an example of outright fraudulent approval, consider DNV, which was suspended by the Executive Board (EB) when it was unable to prove that it adhered to proper safety inspection rules, or even that its staff were qualified to do so (2). TÜV-SÜD too was suspended for giving positive validations in spite of additionality concerns and having staff which lacked the proper qualifications to evaluate projects (6). To further explore the nature of fraud within the context of the network, SGS and DNV will be discussed in detail, and TÜV-SÜD will also be examined.

SGS was established in 1878, initially offering agricultural inspection service. Since then, they’ve expanded to (among others) the automotive, chemical, construction, energy and industrial manufacturing industries (7). In the CDM market, some of the projects that SGS is involved with include Mongolian wind power, involving the UK and China, waste heat use in India, a cement plan, involving the Netherlands, China, and Japan, an Indian wind farm, involving Spain and Sweden, and a manure management system in China involving the Netherlands, Belgium, Canada, and Denmark (8).

SGS (verifier) is linked to three different countries: Brazil (NC=85), China (NC=304), and India (NC=556); thus, it is indeed operating within highly constrained networks for which there is a potential for brokerage. It is also linked to all four of the broker validator nodes, DNV, SGS, TÜV-Nord, and TÜV-SÜD. Thus one can see that a broker typically does not only have a high number of connections outside of its modularity class, and links people who otherwise would not be connected to each other, but also tends to specialize in a wide area and form bridges between highly constrained networks.

DNV was founded in Norway in 1864, to inspect and evaluate the condition of merchant vessels. Since then, they’ve greatly diversified into a variety of areas, such as maritime and oil and gas, energy and sustainability, and product assessment (9). It has a well-established relationship with the international community, as one might expect of an important broker; since 1867, it has created about three-hundred offices in one-hundred countries, and employs over ten-thousand people from over eighty-five nations (9).

Like SGS, it is also involved in a variety of projects. For instance, it is involved in nitrous oxide decomposition in South Africa, along with the UK and Switzerland, a wind power project in inner Mongolia, along with China and the UK, a hydro power station in China, involving the Netherlands, a methane recovery project in Mexico, involving the UK and Switzerland, and a project involving energy efficiency measures at a paper production plant in India, involving Switzerland as well (8).

DNV (verifier) is also as well-connected as SGS. Like SGS, it is linked to India, China, and Brazil. Also, it is linked to the validators of DNV, SGS, TÜV-Nord, and TÜV-SÜD. Again, one can see that not only is it connected to itself, but it’s also connected to entities which are notorious for failing to follow through with the EB’s procedures.

TÜV-SÜD (validator), like DNV (verifier) and SGS (verifier), is linked to Brazil, India, and China. Also, it is linked to the high-outscore verifiers BV Cert, DNVe, LQA, SGS, TÜV-Nord, TÜVRhein, and TÜV-SÜD. One can see that it is well-connected to almost all of the high-outscore validators, while it also holds ties in highly constrained host country networks.

That the biggest players in the CDM market—TÜV-SÜD, DNV, and SGS—are all so closely connected leads to a further understanding of fraudulent activity within the CDM.
only are each of these three entities linked to each other in the validation and verification processes, but they are also linked to themselves. This structure facilitates fraud. A project fraudulently validated by DNV may also be fraudulently verified by SGS, simply because the top players in this oligopolistic market are united in the same scheme. Even clearer yet, a project fraudulently validated by TÜV-SUD may be fraudulently verified by itself later on. Further evidence for this is that when DNV was suspended, “a ‘gentlemen’s’ agreement was installed to help prevent other DOEs from poaching clients from DNV” (2). Overall, the model and gathered evidence suggests that the CDM market is driven by a few validators and verifiers, which in turn are also collaborating with each other in fraudulent activity.

This has significant implications for the success of the CDM as an emissions-reduction program. According to a 2010 rating of DOEs by WWF, TÜV-Nord had the best performance with a D rating, while DNV, SGS, and TÜV-SUD received E+ (3). To determine the ratings, the EB evaluates the projects the DOE positively validated and submitted for registration; the percent of those projects that the EB registers, rejects, and for which it requests corrections determines the rating. For a D rating, only half of the projects positively approved by the DOE are also approved by the EB, 25% are approved after corrections, 20% are registered after a review and corrections, and 5% are rejected (3). For an E rating, only 35% of projects approved by the DOE are registered by the EB; 40% are registered after corrections, 15% are registered after a review and corrections, and 10% are rejected (3). That all the DOEs are given low ratings by WWF lends further credence to the assertion that the CDM market is driven by DOEs united in fraudulent activity. One can conclude that the market is flooded with projects which lack both additionality and sustainability, which seriously compromises the emissions reductions capability of the CDM.

Conclusion

There are ten brokers in the CDM network identified according to the network constraint and outscore parameters. These brokers tend to have broad connections across many markets, work with a wide variety of entities on projects, are linked to other brokers, and operate within highly constrained networks. That this brokerage is at all positive, in that the broker facilitates the flow of good project ideas and forms helpful connections between entities that would otherwise not be connected, is unlikely. Instead, it appears that the market brokers are in fact several DOEs which form an oligopolistic control of project validation and verification. Evidence from this study suggests that these few DOEs are also collaborating with each other in order to further their fraudulent and profitable activities. As a result, the market is flooded with projects which are neither sustainable nor additional; the EB has recognized this by suspending three of the entities most guilty of fraud, DNV, SGS, and TÜV-SUD, and WWF has recognized this in the low ratings it has given to all the DOEs it assessed.

In spite of the fraudulent activity hampering the ability of the CDM to reduce emissions, evidence suggests that there is great room for brokerage within this market. The most productive host countries, China and India, are also some of the most constrained ones; many of the projects operating within their bounds have participants which primarily possess local connections and knowledge. This suggests that these networks have many structural holes within them that could be bridged with brokerage, and this brokerage is currently not occurring because of the heavy presence of fraud in the CDM market. If brokers facilitated connections between members of these networks and those of other groups, and put forth truly additional and sustainable projects, then the CDM would be likely far more effective as an agent for a greener future.

CONTACT MATTHEW JIN AT MATTW.G.JIN.17@DARTMOUTH.EDU

References

Perceptions & Use of BitTorrent P2P File Sharing by Dartmouth College Students

ALEX GERSTEIN a, SCOTT GLADSTONE a,1

Abstract

We investigated student perceptions and use of the BitTorrent peer-to-peer (P2P) file-sharing protocol at Dartmouth College. After providing a general overview of the BitTorrent protocol, its legal implications, and security flaws present, we surveyed students and spoke to Adam Goldstein at Dartmouth College Computing Services in order to better understand the disparity between the perception and actual use of torrenting by students. Using the results from the survey and network bandwidth data obtained from Computing Services, we observed several notable sources of disconnect between student perceptions and incidences of torrenting. These disparities include a perception by students that the average Dartmouth student torrents more than the student participants actually reported, the observation that the number of DMCA complaints far understates the number of students who illegally torrent, and the fact that BitTorrent is still the most utilized file-sharing service at Dartmouth College despite the rise of alternative forms of media management such as video- and audio-streaming services like Netflix and Spotify. This paper outlines the goals, methods, and results of our study and discusses potential policy recommendations for the College in order to curb illegal torrenting and reduce the vulnerability of the Dartmouth College network to torrent-related exploits.

1.1 Introduction

BitTorrent is a protocol supporting peer-to-peer (P2P) file sharing that facilitates the distribution of large data files over the Internet (1). BitTorrent is one of the most common protocols for transferring large files and is often used to distribute popular files and files available for free, such as literary texts, audio files, movies, and applications. “Torrenting” is the process by which a user engages with a BitTorrent client in order to send and receive desired files. To understand the uniqueness and security implications of the BitTorrent protocol, one must first develop a working understanding of the BitTorrent protocol. From there, the legality of P2P file sharing and torrenting become clear. With this foundation, the authors then describe a study in which student perceptions of torrenting at Dartmouth College were compared with College network data on student torrenting. Conclusions and policy recommendations are then made with the goal of aligning student perceptions and administrative desires with regard to torrenting. To the best of the authors’ knowledge, this is the first study conducted that compared student perceptions and actual incidences of torrenting at Dartmouth College.

1.2 Peer-to-Peer (P2P) File Sharing

P2P is a “type of Internet network that allows users with the same program to connect with each other and access files on one another’s hard drives without the intervention of a server computer,” (2). This decentralized network architecture consists of individual nodes called peers that act as both suppliers and consumers of resources. In P2P networks, tasks are shared among multiple, interconnected peers who each make a portion of their resources -- processing power, disk storage, or network bandwidth -- available to all network participants, without the centralized coordination of a server-client model (3).

P2P networks are often found in residential home networks, allowing users to configure their computers to share files, printers, and other resources among all devices; with a number of computers running similar network protocols, P2P is a convenient way to access shared resources. However, the term “P2P” used today generally does not refer to the network architecture from which its name is derived, but to P2P file sharing systems. Using P2P software applications such as Gnutella or eDonkey, users are able to search for, transfer, and download data files over the Internet with any user on the same P2P file-sharing network. The implications of such systems are huge: if any user on the system has a file that another user demands, the second user can - via the network -- obtain the file. The second user will then also become a source of the file in the P2P network, making it faster and easier for more users to download the file.

The first P2P file sharing networks, such as the original iteration of the MP3-sharing network Napster, relied on a central index server to assist with the transfer of files. When someone searched for a file, the central index server -- which contained an index of all of the users and their shared content -- searched for all available copies of the file and presented them to the user; the file would then be transferred directly between the two private computers (4). Because the file sharing occurred over a central network, Napster was held liable for copyright infringement over the sharing of MP3 files and was shut down in 2001. New protocols such as BitTorrent represent a technological evolution of P2P file-sharing networks.
1.3 BitTorrent Protocol

BitTorrent is a P2P file-sharing protocol that makes use of a unique design to reduce the bandwidth and download time for more frequently requested data. The protocol allows users to download a file quickly by integrating all users requesting that file into a “swarm” of hosts to download and upload from each other simultaneously. Thus, instead of attempting to transfer a large file from one user to another -- a slow, bandwidth-intensive process --, BitTorrent allows a user to download small segments of the file from a variety of users, significantly reducing computational energy and download time (1). Because of this design, BitTorrent is often used for distribution of very large files, very popular files, and files available for free, as it is more efficient to distribute these files using BitTorrent than a regular download.

When a user “torrents,” or attempts to download a file over a P2P network using a BitTorrent client, they begin by loading a .torrent file into the BitTorrent client. A .torrent file is a computer file that contains metadata about the files to be distributed, including file names, sizes, checksums of all individual pieces, and a list of the network locations of “trackers.” Trackers are servers that facilitate communication between peers seeking to download the same data on a P2P network (5). The tracker shares the peers’ IP addresses with other BitTorrent clients in the swarm, the group of all peers sharing a torrent. Each piece of a file specified in a torrent is protected by a cryptographic hash contained in the .torrent file (1). This ensures that any modification of the piece can be easily detected, preventing a user from downloading both accidental and malicious modifications of any pieces of a file.

The process of sharing data via BitTorrent starts with a seed, or the initial machine possessing 100% of the data. Then leechers, also known as downloaders, load a .torrent file and request a download. Once connected, a BitTorrent client downloads small pieces of the files specified in the torrent from different computers in the swarm. When the BitTorrent client has some data, it can then begin to upload that data to other BitTorrent clients. By using both downloading and uploading bandwidth simultaneously and dividing the files into pieces, the time required to download the file is significantly reduced (1). Thus, it is worth noting that, "Everyone downloading a torrent is also uploading the same torrent. If 10,000 people are downloading the same file, it doesn’t put a lot of stress on [any] central server. Instead, each downloader contributes upload bandwidth to other downloaders, ensuring the torrent stays fast," (6). As the number of downloaders increases, the reliance of new leechers on old sources decreases. Thus, each successful downloader becomes a source of the file for future downloaders (5). From this point, the process repeats, allowing peers on the network to share any and all files to which any other peer has access.

2.1 Legal Issues & Security Implications

While the BitTorrent protocol itself is legal, those who use the service often do so to obtain files illegally. Illegally obtaining files, also known as digital piracy, involves sharing copyrighted media such as games, music, movies, TV shows, and software that the user does not have permission to share. Whether the user is downloading or uploading content, involvement in this type of non-permitted operation is considered a violation of copyright law (7). Often times, digital piracy constitutes a violation of the Digital Millennium Copyright Act (DMCA). The DMCA is a United States copyright law that criminalizes the production and dissemination of technology, devices, or services intended to circumvent digital rights management (DRM) measures that control access to copyrighted works (8). Penalties for violation include both civil and criminal remedies, which can consist of both significant fines and/or imprisonment. In a 2010 sampling of the one thousand most actively seeded torrent files, 89 percent of the files were confirmed to be illegally shared and the majority of the remaining 11 percent were determined to be “likely infringing,” (9). Only three files, or 0.3%, were confirmed to be legal. The analysis of student’s perception of torrenting in this study will focus on the sharing of the illegal torrents, which comprise the overwhelming majority of the data shared through BitTorrent clients.

2.2 Security Implications for Individuals and Network Hosts

Whenever a user downloads a file, the downloading device is exposed to major security vulnerabilities. With the BitTorrent protocol, where one torrent often contains links to multiple files, users must be cautious about what they are storing onto their hard drives. The distributed nature of torrents allows for them to be intentionally corrupted, either by anti-piracy groups or malicious attackers.

The entertainment industry and various anti-piracy groups have made numerous attempts to thwart those who torrent illegally. These methods, referred to collectively as torrent poisoning, are effective in both slowing download speeds and gathering IP addresses of the downloaders (10). Organizations pose as seeders of a torrent to connect to leechers, or downloaders. Once this connection is made, the anti-piracy organization can save the user’s IP address and use various exploits to deter the user from downloading or sharing the file. Counterpiracy companies, such as the now defunct MediaDefender, slow the download of files through numerous methods. These methods are mostly based on seeding or linking to fake files. For example, in a decoy insertion, corrupted versions of a file are distributed on a network to appear indistinguishable from the original files. Attempting to download these files often leads to excessive download times and potential insertion of malware into a user’s computer. While a decoy insertion targets the torrent files, a spoofing attack targets the locations of torrents to redirect potential leechers to non-existent locations. Spoofing attacks aim at consuming all of a BitTorrent client’s bandwidth resources by having the client to try to grab a file that doesn’t exist, thus preventing the user from downloading or uploading any other files.

Both these attacks are founded on the principle that some downloaders will stop torrenting out of frustration with slow download speeds and corrupted files. However, sometimes copyright enforcers take extreme measures to proactively prevent BitTorrenting. In 2007, Comcast was accused of hindering the uploading of complete files to BitTorrent (11). In an effort to increase network access speeds, Comcast blocked torrenting applications by using their routers to monitor and prevent certain forms of network traffic. The following year, they were ordered to terminate their "unreasonable" network management. In another instance, MediaDefender overstepped their legal rights similarly in 2008 when they caused a Denial
of Service attack on Revision3, a legal Internet television network distributor (12).

However, it is not only the anti-piracy organizations that corrupt torrents. According to multiple studies, BitTorrent is among the most frequently used mechanisms for attackers to distribute malware. One report claims that up to 14.5% of BitTorrent downloads contain malware and 47% of all zero-day malware found was distributed through BitTorrent (13). With these exploits, attackers are able to execute shell code, install keyloggers, and create zombies on unsuspecting users’ computers.

### 3.2.3 Perception Data

The participants in the survey were 130 students at Dartmouth College. The population consisted of second-year, third-year, fourth-year, and graduate students at the College. Participation was anonymous and voluntary. All prospective participants were sent a link to a Google Form that contained all of the survey questions. All participants had completely voluntary choice in clicking on the link, examining the questions, and completing and submitting the survey. The population examined in this study may be biased since some individuals who completed the survey may have personally known the researchers or were members of CS 55: Security & Privacy, the class for which this research was completed.

#### 3.2.2 Survey Design

The survey consisted of seven questions, six of which were mandatory. The mandatory questions asked the participant for his/her personal experience with torrenting, whether he/she believed other students at Dartmouth torrented, the number of files that he/she believed the average student torrented, whether he/she believed Dartmouth recorded information about student torrenting, and whether his/her knowledge of Dartmouth Copyright Violation Policy impacted the amount that he/she torrented. The optional seventh question asked whether the participant had any thoughts as to how Dartmouth College could reduce the incidence of torrenting on campus. All surveys were timestamped when submitted and resubmissions/changing of answers were not permitted.

#### 3.2.3 Perception Data

Of the 130 students, 78 (60.0%) stated that they had previously downloaded a file via torrenting. 52 out of 78 individuals (74.3%) who had torrented in general had also torrented while at Dartmouth. Of the 52 students who designated that they torrented at Dartmouth, 34 (65.4%) claimed to download one to five files per month on average; 5 (9.6%) claimed to download zero files per month on average; 5 (9.6%) claimed to download six to ten files per month on average; and 9 (17.3%) claimed to download 11 or more files per month on average.

Considering all 130 students (both torrenters and non-torrenters), 82 (63.1%) claimed to download zero files per month on average via torrenting. This was statistically significantly different (p<.00001) from the small percentage of students (16 students, 12.3%) who reported that they believed that the average Dartmouth student downloads zero files per month via torrenting. The median self-reported number of files downloaded via torrenting on average was zero, while the median perceived number of files torrented by the average student was in the category “one to five.” Based on the distributions presented below, there is a clear right shift in perceived files downloaded as compared with reported files downloaded, indicating that there is likely a significantly higher perception of student torrenting at Dartmouth than what actually occurs.

The results were less conclusive with regard to whether participants believed that Dartmouth recorded student torrenting information and whether knowledge of Dartmouth’s Copyright Violation Policy affects how often the participant torrented. Of the 130 students asked if Dartmouth records when someone torrent a file, 47 (36.2%) said Yes; 28 (21.5%) said No; and 55 (42.3%) said they were Unsure. Attempting to break down the results further by the results of other questions in the survey, the results were still inconclusive. More data would need to be collected to determine this result. Regarding
whether knowledge of Dartmouth’s Copyright Violation Policy affected personal torrenting practices, there was no significant difference \((p > .05)\) between the number of participants who said Yes \((N=63, 48.5\%)\) and who said No \((N=67, 51.5\%)\).

### 3.3 Dartmouth Network Patterns

The researchers in this study met with Adam Goldstein, IT Security Engineer at Dartmouth College Computing Services, to discuss Dartmouth network trends and use of BitTorrent and other file sharing protocols on the Dartmouth network (16). Goldstein described the PaloAlto firewall and traffic recording system utilized by Dartmouth, provided information on gigabyte usage by the top file sharing applications used on the Dartmouth network, and discussed the incidences of Digital Millennium Copyright Act (DMCA) complaints and violations on campus.

#### 3.3.1 PaloAlto System

PaloAlto is a layer-7 firewall employed by Dartmouth College to prevent proliferation of malware and provide general security protection for the Dartmouth College network. Dartmouth currently uses the system for protection purposes only, but the PaloAlto system is designed such that it could be used for traffic monitoring, controlling network access, and designating bandwidth rules. However, Goldstein made clear that Dartmouth does not currently use the system to police on network traffic or content.

The PaloAlto system has a variety of features that are worth noting. First, the system includes application identification, or “App-ID.” By classifying traffic based on application, instead of exclusively by port, system administrators can note the types of applications commonly being used by network users. Additionally, the system provides “User-ID” and “Content-ID” information, allowing network administrators full visibility and control over network traffic. Finally, the system provides detailed sandbox analysis -- known as “WildFire” -- of suspect incoming executables. WildFire allows the network to adapt and learn about new types of malware without being explicitly provided with a list of known targets. The system accomplishes these goals by performing network analysis on the entirety of all network packets. With the massive amount of data, the system is able to adapt and evolve based on current configurations and usage trends (17).

#### 3.3.2 BitTorrent Usage and File Sharing Application Data Consumption

According to Dartmouth College Computing Services, user traffic to the Internet averages about 3.5 Terrabytes per day, with BitTorrent occupying approximately 4.8% of total network traffic (16). While BitTorrent usage has declined over the last decade and has been replaced in a large capacity by audio- and video-streaming applications such as Netflix and Spotify, BitTorrent is still the most highly used file sharing protocol at Dartmouth College (18). Goldstein noted that there has been a significant decrease in the use of previously highly utilized file-hosting service MegaUpload; this is not surprising as this service was shut down in early 2012 by the United States Department of Justice.

Referring to the file-sharing application consumption data presented in the set of figures entitled “File Sharing Gigabyte Usage by Application,” BitTorrent comprises 70.5% \((177.09\ GB)\) of file-sharing traffic in a 24-Hour period, 60.1% \((1180.17\ GB)\) in a 7-Day period, and 60.9% \((4757.79\ GB)\) in a 30-Day Period, all ending November 8, 2013. Goldstein attributed this high level of data consumption to the downloading of large video files and software applications; he noted that students are often identified by DMCA complaints for torrenting full seasons of television series, recently released films, and commonly-utilized software applications such as the Adobe CS Workshop.

Regarding the incidence of security concerns caused by BitTorrent usage at Dartmouth, Goldstein cited the large percentage of Apple computers on campus as a force providing significant protection to many network users. The amount of malware targeting Macs is significantly lower than that targeting PCs, said Goldstein, due to the lower market share occupied by Apple computers. However, network security concerns caused by torrents are continually present, including an issue this past year in early 2013 when malware was distributed through the network via fake textbook torrents. The widespread proliferation of the malware was contained by the PaloAlto system, which identified the malware through its “WildFire” analysis.

#### 3.3.3 DMCA Complaints at Dartmouth College

Each year, Dartmouth College Computing Services receives a number of complaints of violations of the Digital Millennium Copyright Act (DMCA) that have occurred on the Dartmouth network. Current policy dictates that, upon receiving the complaint from an authorized organization, Dartmouth is required to send a copy of that complaint on to the individual responsible and request that all incidences of the downloaded files be removed from the individual’s computer. If a student receives multiple complaints, usually on the order of 3 or more according to Goldstein, they are referred by Computing Services to Dartmouth College Judicial Affairs for disciplinary action. In 2012, Dartmouth received 466 DMCA complaints. Through October 2013, Dartmouth has received 343
complaints, which is consistent with the number of complaints received last year. Goldstein expects the number of complaints to remain constant in 2013.

Goldstein noted a number of problems with receiving DMCA complaints and attempts to connect the complaints with those individuals responsible. According to Goldstein, Dartmouth’s wireless network is split into six zones; thus, it is possible that one person illegally sharing a season of a television show can generate 50-60 complaints due to address translation and physical movement of his/her laptop around campus. Additionally, a complaint often cannot identify a specific individual’s IP address, but that of a router or a shared access point. In these situations, College policy dictates that the person to whom the router or access point is registered is responsible for the DMCA complaint. This tends to lead to false identifications of individuals who have registered routers but never used a BitTorrent client, as it is possible that other users utilizing the router are the cause of the complaint. Lastly, Goldstein made light of the fact that many individuals are able to successfully torrent a copyrighted file without ever receiving a DMCA complaint, while other individuals are identified for their first-ever incidence of torrenting on Dartmouth’s network. Goldstein stated that certain files tend to be targeted at different times by different organizations. For example, HBO aggressively targeted torrenters of “Game of Thrones” earlier this year. This discrepancy between downloading files and being caught makes it difficult to curb individual torrenting behavior.

3.4 Relationship between Student Perception and Dartmouth Network Data

There exists an inherent discrepancy between the types of data collected in the different sections of the study: the perception survey targeted individual users’ actions and beliefs, while the network data was aggregated for all students at Dartmouth. While this makes some one-to-one comparisons between actual use and perception of torrenting difficult, there are still some valuable relationships between how students perceive BitTorrent use and its actual incidence on campus.

Generally, the number of students who received DMCA complaints each year is relatively small compared to the size of the student body. The 466 complaints from 2012, for example, is relatively small compared to Dartmouth’s approximately 6000 undergraduate and graduate students. These numbers indicate that a maximum of 7.8% of students receive complaints in a given year. However, as Goldstein pointed out, one student’s actions could generate 50-60 complaints. For this reason, the percentage of students who receive DMCA complaints is likely much smaller, closer to 1-3% of the total student population. This implies that a large portion of students who torrent in a given year never receive a DMCA complaint.

From the data on whether students believe that Dartmouth records when someone torrents a file, it is clear that the 21.5% of students who answered No and the 42% who answered Unsure underestimate the amount of information to which Dartmouth’s network administrators have access. There are a number of possibilities for why 63.5% of students are not aware of Dartmouth’s access to their internet behavior. Students might not consider Dartmouth’s access to their personal information because they prefer to turn a blind eye to the associated privacy issues. Perhaps they believe that Dartmouth does not have the technical capabilities to keep this information. Nevertheless, the fact that the No/Unsure segment is so large indicates that there is a clear disconnect between students’ perception of the Dartmouth’s network capabilities and the actual practice by the College. It is possible that, by equipping students with knowledge of Dartmouth’s capabilities, student network usage and Dartmouth practice may become more aligned.

4.1 Conclusions and Policy Recommendations

Through the research, students’ torrenting patterns and perceptions were measured via self-reported survey data and Dartmouth College Computing Services provided data on network bandwidth usage and information on DMCA complaints. Torrenting activities continue to take up a non-negligible percentage of network bandwidth, yet only a very small fraction of these students receive DMCA complaints that might curb their illegal activities. If students, administrators, and copyright enforcers aligned their interests, then perhaps network security would improve and copyright infringement
would decline.

Although BitTorrent use is already subsiding with the rise of online media streaming, there are some policy changes that might expedite the decline. One clear way to decrease illegal torrenting of files would be for the College to provide free, alternative access to current media. With the rise of streaming services juxtaposing the slow fall of BitTorrent, it should be evident that convenience is a major factor in determining how a user watches a television show or listens to music. If there were a way for Dartmouth to acquire a corporate license for legal streaming services such as Netflix, HBOGO, and Spotify, then the amount of students obtaining their media legally would likely increase. This suggestion was brought up by 23.5% of survey participants who provided possible recommendations. Although the amount of bandwidth required by BitTorrent downloads is not drastically different than the streaming alternatives, providing these legal streaming services would be able decrease DMCA complaints and the amount of malware that can pass through Dartmouth’s networks.

One of the other most common recommendations by survey participants, was to increase the disciplinary action and/or consequences of illegal torrenting. 14.7% of students stated that, instead of waiting for the anti-piracy firms of major movie and music studios to reach out to students, the College can take the initiative of preemptive action. This approach would work by drastically reducing the population of students (63.5%) who did not believe or were not aware of what Dartmouth College Computing Services is capable of tracking. If students know this information, perhaps they would reduce their usage of BitTorrent.

Tied with increased disciplinary action among student recommendations was to provide discounts on textbooks and software. The Dartmouth Computing Store already provides many student discounts on software for students, but they might not advertise their discounts in the best way possible. Additionally, consumers would still prefer free to paying, so this course of action may not yield the best results.

Despite the useful recommendations among those who answered the optional final question, nearly half (47.1%) believed that no changes were necessary. These people might not want any changes because either they want to continue their illegal actions or they don’t torrent anyway. With BitTorrent usage declining without further preventative measures, this group might have a point that no further action is completely necessary.

However, the discrepancy between student perceptions and College actions may extend to how students think about security in a more general sense. If users never understand the consequences of their actions, then they will see no reason to change their online behavior. Through taking more severe measures against students’ risky behavior on the network, the administration might be able to help curb security concerns.

Acknowledgments

Charles Palmer provided guidance in the direction of this research and helped to connect the authors with security specialists at Dartmouth College Computing Services. Adam Goldstein from Dartmouth College Computing Services engaged in useful discussion with the authors and provided key data about Dartmouth network patterns and BitTorrent usage, without which this research could not have been completed.

References


CONTACT ALEX GERSTEIN AT ALEXANDER.S.GERSTEIN.15@DARTMOUTH.EDU

CONTACT SCOTT GLADSTONE AT SCOTT.W.GLADSTONE.15@DARTMOUTH.EDU
Effects of Habitat Disturbance on the Parasites of *Propithecus edwardsi*

**A. CHASE SCHOELKOPF***

Abstract

Parasitology can provide invaluable knowledge for the conservation of flora and fauna. To truly understand the health of a species and the condition of its habitat, one must consider the parasites present. This study examines the parasites of *Propithecus edwardsi* in Ranomafana National Park, Madagascar, and the effects of habitat disturbance on parasite richness. Three previously unreported parasites were identified in *P. edwardsi* through noninvasive fecal sampling; however, a comparison of parasite richness between undisturbed forest and disturbed forest did not yield a statistically significant difference. These parasites suggest cross-species transmission and the need for follow-up studies to track the movement of these parasites.

Introduction

Madagascar and its endemic lemurs remain top conservation priorities, as deforestation and other human impacts continue to threaten their survival. Between 1950 and 2000, nearly 80 percent of Madagascar’s forests have disappeared (1). The disappearance of forests primarily impacts lemurs directly by destroying their habitat, but can also affect lemurs indirectly in a number of ways.

One indirect manner is habitat disturbance’s effect on lemur parasitology and, in turn, lemur health. Animals become infected with parasites by directly contacting parasites, ingesting parasites, or ingesting an intermediate host (2). As a habitat becomes disturbed, these methods of infection become more likely. When the habitat of a species is disturbed, the species has to adapt to the reduction in resources. These adaptations often take the form of longer travel distances and the exploration of new territory. Longer travel distances have been shown to have a direct correlation with higher parasite prevalence (3), while travelling into new territory sets up the possibility of cross-species contamination. While searching for food, individuals are more likely to travel into areas containing humans and their domesticated animals (4). Of particular concern in Madagascar is the transmission of parasites from zebu (*Bos indicus*) and pigs (*Sus scrofa domesticus*) to lemur species (5). In addition to contact with domesticated animals, traveling into new areas also increases the contact between humans and primates and, as a result, the possibility of transmission between the two (6).

With an increase in habitat disturbance, an increase in fleas and *Rattus rattus* also occurs (7, 8). Fleas can act as intermediates for a variety of pathogens and therefore could transmit a parasite directly to a primate host or could transmit another disease that weakens the primate. In a weakened state, a primate’s immune system will not function at high levels and the host will have a lower likelihood of defending itself from parasitic infection. *R. Rattus* in Madagascar has been shown to contain up to 12 different parasites in addition to the bubonic plague (8). Therefore, *R. Rattus* presents a large risk for direct transmission of parasites to primates as well as weakening immune systems through other diseases.

Stress can also play a large role in the relationship between habitat disturbance and parasitic infection. Habitat disturbance can lead to higher competition for food, an increase in aggressive encounters, and more expensive foraging costs, all of which can elevate stress levels and overtax the immune system (9-11). With a depressed immune system, any parasitic infection may be augmented (12-17).

*Propithecus edwardsi*, the Milne-Edwards’ Sifaka, is a large folivorous-frugivorous lemur species found in the southeastern rainforests of Madagascar (18). *P. edwardsi* is listed as an endangered species on the IUCN Red List and has been the subject of long term study for over 25 years in Ranomafana National Park, RNP (19). RNP is a submontane rainforest in the southeast of Madagascar that has been a national park since 1991. However,
prior to the park’s establishment, parts of the forest were subject to heavy logging. The parasites of *P. edwardsi* were studied in 2001, but no further research has been performed on the topic since that time. In 2001, the *P. edwardsi* population of RNP was found to have six different classes of endoparasites (20). Here the gastro-intestinal parasites of four groups of *P. edwardsi* are identified to examine the effect of habitat disturbance on parasite richness.

**Methods**

Fecal samples were collected just before the rainy season, between October 29 and November 12, 2013, within Ranomafana National Park (RNP) in Southeastern Madagascar (21°16’S, 47°20’E), (19). Ranomafana contains two separate sites, Valohoaka and Talatakely, with habituated *P. edwardsi* groups (Figure 1). Samples were collected from 4 individuals in Talatakely, a disturbed secondary forest, and from 10 individuals in Valohoaka, an undisturbed primary forest. Talatakely was selectively logged as recently as 25 years ago and hosts the vast majority of RNP’s 27,000 tourists per year (19, 21). Meanwhile, Valohoaka is an area around 8.5km southeast of Talatakely that has never been selectively logged and receives minimal tourists each year (22).

Fecal samples were opportunistically collected from each individual and were stored in two ways. First, a portion of the sample was stored in 10 percent neutral-buffered formalin and labeled. A second portion was stored fresh in aluminum foil and labeled. The samples were then transported back to the laboratory at Centre ValBio for analysis. The samples preserved in formalin were used in a modified Fülleborn flotation technique to reveal the presence of helminth eggs and larvae. For each individual, a portion of the sample was placed into two Eppendorf tubes. These tubes were centrifuged for two minutes, and the supernatants were removed. A 20 percent sodium chloride flotation solution was vigorously mixed with the pellets and allowed to sit for 45 minutes. After the flotation, the top layers of fluid from each Eppendorf tube were placed onto four microscope slides, for a total of eight microscope slides per individual. The slides were then read at 10x magnification using a compact light microscope.

The fresh sample was used in a modified Baermann sedimentation technique. The sample was wrapped in a Kimtech wipe and placed in a funnel. The funnel was clipped at the bottom and filled with water until half of the wrapped sample was submerged. The sample sat for a two-day period, allowing the eggs to hatch into larvae. After two days, the solution was drained into an Eppendorf tube and centrifuged for 10 minutes. The supernatant was drawn off except for the bottom 0.2 milliliters (ml). This 0.2ml of sediment was then placed on a compound light microscope slide for reading at 10x magnification. Sedimentation provided a second level of analysis for the helminths and provided a method to reveal the presence of flukes too heavy to float in the Fülleborn technique. After recording the parasite richness in each individual, a Mann-Whitney U-Test was used to test for a significant difference between the individuals of Valohoaka and the individuals of Talatakely. A non-parametric test was needed due to non-normally distributed data and the unequal sample sizes.

**Results**

In the groups of both study sites, three kinds of parasites were identified: *Toxocara* sp., *Ancylostoma* sp., and *Strongyloides* sp. *Toxocara* sp. was identified by the thick wall surrounding a dark, circular embryo in its egg form and by the clear esophagus portion, denser intestine portion, and nub tail of its larva (23, 24). *Ancylostoma* sp. was identified by its oval shape, smooth wall, and clustered grouping of cells inside the egg and by the long tail of its larva (23, 25). *Strongyloides* sp. can be identified by the thin oval shaped walls of its egg and by the nub tail of its larva (23, 25). In Valohoaka, five out of the 10 individuals had *Toxocara* sp., but only one individual had *Ancylostoma* sp. present, and one individual had *Strongyloides* sp. present (Table 1). In Talatakely, one out of four individuals was infected by *Toxocara* sp. Meanwhile, two individuals had *Ancylostoma* sp. present, and two individuals had *Strongyloides* sp. present (Table 1).

In Valohoaka, four out of 10 individuals had zero parasites. Five individuals had one type of parasite, and one individual had...
two types of parasites. In Talatakely, one out of four individuals had zero parasites, one individual had one type of parasite, and two individuals had two types of parasites (Figure 2). No individual in either location had all three types of parasites present. These values led to the individuals in Valohoaka having an average of 0.7 different parasites. In Talatakely, the individuals had an average of 1.25 different parasites (Figure 3). However, the Mann-Whitney U-Test revealed that there was not a significant difference in parasite richness between the two groups with a p-value of 0.3.

Discussion

This study suggests that the *P. edwardsi* individuals in a disturbed habitat do not have a higher parasite richness than those individuals in an undisturbed habitat. There are several possible explanations for this result, each of which should be the subject of future research. There has been recent movement of individuals between Talatakely and Valohoaka that may have led to a more uniform parasite richness between the locations. Parasites lay eggs on a monthly cycle, so there is a possibility that the period of study did not include the gestation of these parasites. Unfortunately, the study had a very small sample size. The *P. edwardsi* population in Valohoaka and Talatakely has been reduced to only four habituated groups. Compounding this issue was the fact that the second group in Talatakely was unable to be located in order to collect samples.

Interesting results occurred when comparing the current parasitology to the parasitology in 2001. In this study, three parasites were observed that had never previously been reported in *P. edwardsi*: *Toxocara* sp., *Ancylostoma* sp., and *Strongyloides* sp. Previously, *Lemuricola* sp., *Monezia* sp., *Anoplocephala* sp., *Strongylus* sp., and *Physocyclus* sp. were encountered (20). Parasites lay eggs on a monthly cycle, so there is a possibility that the period of this study did not include the gestation of the parasites previously found. The second possibility is the shifting of parasite populations in the last 12 years in Ranomafana National Park.

Moving forward, the scope of this study should be expanded. By lengthening the period of study, there would be a higher likelihood of encountering all parasites present in *P. edwardsi*. Extending the field of study to protists could also be highly interesting. During my study, protists were observed in six out of 14 individuals. However, I did not have the knowledge or tools to attempt to classify these protists.

Although the *P. edwardsi* of Ranomafana appear to be healthy, it is important that their parasitology continues to be monitored. Parasites have been known to lead to wildlife declines, and as an endangered species, *P. edwardsi* cannot afford any such declines. Even researchers who are not studying parasitology should be aware of the clinical signs of parasites such as anemia, continued diarrhea, respiratory issues, or unexplained death. If they noticed any of these dangerous signs, they could inform researchers or veterinarians studying parasitology.

The pathogenicity of these parasites in lemurs is not well known, but clinical indications can be extrapolated from other species. *Toxocara* sp. in dogs and cats can lead to stillbirths, neonatal deaths, stunted growth, chronic ill thrift, diarrhea, and pot-belly appearance. *Ancylostoma* sp. in dogs and cats can cause anemia, poor hair coat, death, and diarrhea. *Strongyloides* sp. in pigs, zebu, dogs, and humans can result in diarrhea, dehydration, respiratory issues and death (23, 26).

Continued monitoring is also necessary to track the spread of these parasites to other species inside Ranomafana National Park. Finally, further research needs to be performed in order to determine the original host for these parasites. By pinpointing the source of these parasites, perhaps steps can be taken to prevent their spread and to protect the unique fauna of Madagascar.

Acknowledgements

I would like to thank the Institute for the Conservation of Tropical Environments, Dr. Tharcisse Ukinamitambara, and Franck Rabehary for their logistical and technical support. Collection of fecal samples was greatly aided by my guides Radrianasolo Laurent, Randrianirina Georges, Rakotovao Remi, and Rakotjonjavoto Justin. I extend a special thanks to Rafalinirina Herman Andry for help with methods and parasite identification. Dr. Clara Scarry assisted with the statistical tests. Finally, I thank Tom and Patty Schoelkopf for providing the funding to make this research possible.
References

8. Lehtonen J. Occurrence of the introduced black rat (Rattus rattus) and its potential effects on endemic rodents in southeastern Madagascar. Helsinki, Finland: University of Helsinki. Ph.D.
25. Rafalirina Herman Andry, personal communication.