Speaker	Professor Michael Small
Talk title	Why your friends have more friends than you
Venue	Helvetica
Time	Tuesday 29 October 2019, 5:30pm

Simon Handford

Good Evening everybody.

Thanks for coming, it is really pleasing to see so many people coming to Raising the Bar this evening. My name is Simon Handford and I am from the Research Commercialisation Office at the University of Western Australia and I will be introducing our speakers tonight.

Before we kick off though I would really like to just acknowledge the Whadjuk Noongar people as the traditional owners of the lands and waters where we are meeting today and to pay my respects to the elders, past, present and emerging.

This evening is all about sharing some of the amazing research that happens at the University of Western Australia with the general public and the community at large, so welcome to Raising the Bar tonight.

We have got twenty researchers and academics speaking across ten bars in Perth all around the CBD hopefully with excited audiences like we have got here tonight.

At UWA we are excited to make education a part of our cities popular culture, through transforming local city bars into a place where you can enjoy a drink, whilst learning about the impact that our research has in the community at large.

If you are sharing Raising the Bar on social media we would love to get Raising the Bar trending tonight, so the hashtags are @UWAresearch or #rtbperth19. A bit of tweeting would be appreciated to spread the word and just to let you know that tonight's talks are being recorded and will be published as podcasts on our social media channels in a few weeks' time.

There might also be a few people taking photographs around and tweeting photos and the likes, so I hope that is okay with everybody.

We have got two speakers this evening and I am going to be introducing Professor Michael Small, first.

Michael is an Applied Mathematician and holds a CSIRO UWA Chair of Complex Engineering Systems. Michael's research is in the area of Complex Systems, Chaos and Dynamical Systems Theory. He uses these methods to investigate a range of problems from the spread of disease and effective play and

football teams to the flocking of birds, emergent machine learning and predictive failure of industrial machine components and I am told that he has practised this on his eight-year-old. So, I am really hoping that is going to help us tonight.

Michael, you have got the stage, you have got a microphone, you have got no props and you have got half an hour to talk about why … now, what are we talking about? We are talking about why you have more friends than I do. Oh, they've got more friends than you! Okay. I shall hand over the mic.

Please, Michael Small

[clapping]

Professor Michael Small

Thanks a lot and thank you, so many of you for turning up tonight, on a Tuesday night for what I hope your realised was going to be a maths talk [laughing].

So, why your friends have more friends than you do.

This is called "the friendship paradox" and I will get to explaining it in detail, a little bit of detail later on. Hopefully, I will keep it understandable. If it doesn't make sense, please just yell out and tell me that I am just rambling to myself and so let me start by just describing what I mean by the problem and then I will talk a bit about the research that we do and then I will get back to the problem at the end.

What the friendship paradox says, is if you look at your social network, either if you are on Twitter or Facebook or Instagram or whatever, then the number of friends that you have and the number of friends that your friends have, on average, your friends will have lots more friends than you.

It is a good news story, so you don't have to feel inadequate if all your friends have got more friends than you.

It is also true, to a certain extent, it is also true for real networks, so your actual friends on average, your friends are more popular than you but the good news is, it is a mathematical fact and it is not just that there is something wrong with you.

I don't know you [laughing] but the mathematics says the definitely your friends have more friends than you. My research, there were lots of big long words and that but technically it is in the area of Complex Systems and Chaos and what that means really, I am interested in two sorts of problems.

One, where we want to predict what is going to happen in the future and we measure what is happening now and we try and use that to build a model which will tell us what is going to happen next, and the other, there are complex systems, is all about understanding how people or individual components interact. The thing we should be thinking about is a network, where you have lots of nodes or a social network, you and all your friends are the nodes on that network and there are

connections between them and that forms a network, so that is kind of the mental image you have to have for the rest of the talk.

Does that make sense?

Yes?

So just a bunch of nodes, a bunch of points connected by lines and each of those nodes is you and one of your friends and the lines there, tell you if you are friends so you are connected.

You have got the lines connecting you to your friends. But first of all, I am told I have to talk a bit about my own research and what I am doing.

I am involved with three fairly large projects at the moment. One of them, the clue was in the title that CSIRO UWA blah, blah, blah, blah. That means that UWA only pays half my salary, the other half comes from CSIRO through their Mineral Resources Unit and through that, I am looking at applying mathematics techniques to problems in the resources industry.

These are largely problems about finding stuff underground, figuring how to dig it up, process it, so you don't make too much of a mess of the environment and put back all the bits that you don't want and ship off the good stuff to make money.

I am not going to talk about that.

The second project that I am involved with, we have a research centre where we are looking at applying data science to problems in maintenance and again, this is related to the resources sector and this is in partnership with three other companies up on the Terrace here and essentially, what we are doing is we are looking at all their big, yellow machines up in the Pilbara. They have got lots of instrumentation on them, measuring all sorts of stuff, collecting all sorts of data and what we are trying to do, is use that data to tell them when to maintain the machines because they are going to break down.

The final project that I am involved with is a UWA initiative called "Young Lives Matter" and it is about suicide and mental health in youth and adolescence and mathematically there is a similarity between that and the problem in maintenance that I just described.

We are trying to understand at least an approximate model of mental health and try to use that to predict who and when people are at risk.

So, those are the kind of applications that we have visited and after I explain the friendship paradox, hopefully, I will be able to show how that actually connects to some of them.

So, I need a drink. [laughing]

So, the friendship paradox. On average, your friends have more friends than you, so on average, how many friends do you have? Well, that just depends on how many connections the average person

has in this network, on Facebook or whatever, but the key is to realise that when I say, "your friends" now, your friends are not average people. Well, [laughing] some of my friends are standing up there and they are definitely quite average [laughing]. Your friends are not average people they are average friends, so, if I want to take an average over your friends, I am taking an average over those edges, over people that are connected to you. And the key thing is, if someone has lots of friends, then they are going to appear in that average lots and lots of time.

There is more chance of finding someone with lots of friends because people with lots of friends have lots of friends.

Another way of thinking about it, if I pick an average person in this room. Thomas is extremely average, ask him how many friends he has, he has got one hand free so he can count the total number, but each of those friends, they are not themselves, average random people, they are average random friendships we are choosing. Right? So by choosing one of his friends I am choosing an average friendship and people with lots of friends will be in lots of friendships, so when I take that average there is more chance of finding those people that have lots of friends.

Does that make sense?

Good! My eight year old said it made sense but I promised him chocolate if he agreed. [laughing]

So, that is the basic idea of the friendship paradox and the other thing that makes it work, is it actually depends on the, what we call mathematics, the degree distribution. That is the number of connections on average, the average person has and for certain types of distributions, so for certain average numbers of friends, this property is extreme, and it turns out that this phenomenon is most extreme in exactly the sort of network that you get in social networks and the reason for that is something called preferential attachment.

I am throwing out lots of "buzz" words. If you don't like the buzz words just let them go past you and I will give you the details.

The thing with preferential attachment is, if you want to … if you add someone to a social network, how would they join that network. They would connect to a few people but how do they choose those people to connect with?

On average, if typically, when you join a social network, you are going to try to connect to popular people. Why would you join a social network if Thomas was your only friend? You want to find some popular people to talk to as well. If you join a social network, you are going to connect to … [laughing] How many friends do you have? What was I saying apart from insulting Thomas?

As new people join the network, they will on average try to connect to the most popular or the more popular people. That is just a natural kind of phenomenon. This means that the people that have lots of friends get even more friends. This is what comes up in economics and social sciences as The Burrito Principle or the idea that the rich get richer.

I was reading in the newspaper yesterday that some genius in the West Australian financial column, don't raise your eyebrows like that [laughing] some genius in the West Australian financial column said that Warren Buffett made 99% of his money after he was sixty. Therefore there is hope for all of us to make a hundred billion dollars because he made ninety-nine billion dollars after sixty but as most of you, and if you don't know this, please come and talk to me. As most of you know, earning the first billion dollars is quite hard [laughing] Once you have got lots of money it is easy to make more money. If you start with a small loan of a million dollars, I told you I would get a Donald Trump joke in somewhere! If you start with a small loan of a million dollars it is easy to make some money. If you don't have that money to start with, it's much harder.

So what this means is, in social networks, people with lots of friends get lots more friends but in economic systems, wealth is centralised. Eighty per cent of the wealth was with twenty per cent of the people, for example. That is a general principle that follows in all of these types of systems.

I have told you why your friends have more friends than you do and why that is not cause to be depressed unless your only friend is Thomas, then, in which case …..

I have told you how this comes about in social networks. It comes about from this preferential attachment. This idea as people join the network, they connect to the most popular people, but how can we really use this. I mean, Twitter is all very well but it is not really all that important to most of us.

Let me give you a couple of examples and then I will finish and hopefully, I am not too quick or too long. I will give you a couple of examples where we actually apply this sort of thing.

We started looking at this and applying it to problems of disease transmission. Back in 2003 I was working in a university in Hong Kong and the Sars outbreak came along, so an interesting time to be in the city and as a university researcher you are always looking for research funds and when you have got something like this on your doorstep it is clear that this is an important problem and a worthy focus of research because, apart from the fact it might help people, someone might actually pay you to do the research.

We were looking at disease transmission and you can transmit diseases like information in a social network. When Thomas starts tweeting about how fabulous this talk is, that spreads through his social network to his friends and his friend's friends and so on and eventually reaches someone with more than three friends and keeps going but in real networks, diseases are spread through contact. People that you are close to.

The basic model for this is no different. On average, you have a certain number of connections that you see on a daily basis, friends, family, work colleagues. You have got some other number of people that you happen to bump into at a bar, on a bus or whatever and depending on the type of disease that you are modelling, those are the people that are going to get infected from you.

We were able to show that disease transmission looks a lot like friendship on social networks and in particular, if you think about, I've lost my train of thought. If you think about how you treat a disease? You want to apply immunisation. I have got a bucket of money to get immunisation to everyone, ideally, I immunise everyone in the population but that is expensive. I can't afford to do that, so I have got enough doses of immunisation for maybe ten per cent of the population, so what do I do? I want to immunise the people who are most likely to spread the disease. Who are they? Well, they are the people with the most friends because they have got the most contacts. People like Thomas, he doesn't see many people, so he is probably a low priority but if you have got more friends, then you have got more chance of infecting other people because you have more contact with other people.

How do you find those people?

If I say, "Well I have got a cure for this deadly disease, who wants it?" Everyone says they want it. Everyone thinks they are important. So, how do I find the people that have lots of connections? If I just choose random people, I get someone with an average number of friends but if I choose average people and ask them to nominate one of their friends, I get people with a lot more connections. If I have got limited vaccination, what I can do is go to Thomas and say, "I am not giving it to you. I am not giving you the immunisation but name one of your friends who you think deserves it, I will give it to them." So, Lyle gets the immunisation because Lyle sees a lot more people and he is immunised against spreading the disease and I have eliminated one of the hubs, one of the nodes, one of the people in the network with a lot of connections. So, this simple idea of friendship and connections between people actually has very simple, very direct, real applications in things like disease transmission.

I have to take another drink of beer so that I can try and remember the second example that I was going to talk to you about.

Another example that we looked at is football. By football, I mean real football, Australian Rules Football [laughing] I am a Kiwi, we lost the World Cup so I am going back to AFL now. AFL, the reason AFL is interesting, well there are lots of reason this is interesting, but the important reason is interesting is because it is a game that goes on forever [laughing] it has got a very high score unlike soccer where you are just measuring ones and zeros and it is kind of boring, sorry Thomas. It has got a lot of players on the team and it's got no rules.

The good thing about having no rules is there is no position, there is no off-side, there is no forward passing, there is none of this stuff which gets in the way of the basic idea of getting the ball from there up to there. So, you have got two teams of… I don't know … what is it? Seventy-four people, each trying to move the ball from one end of the ground to the other in opposition to each other. Essentially, those two groups of people are acting like trying to transport something. Like a fluid being transported from one side of a field to another, but you have got these discreet individuals acting in that system and they are connected to the other individuals in their team if they are passing the ball between them

As I said, AFL, there are no rules so by passing, I just mean the ball gets from one person to the next person. It doesn't really matter, as long as you touch the ball and then he's the next person who touches the ball, that's a pass. We took the data from a season of AFL and we looked at the network that we get when we map all the players as "nodes" and the connections between them, they are friends if they are passing the ball from one to another, over a five-minute interval, for example. So, it is a moving window over the game. If you imagine this graph, this network of nodes is changing as you go through the game and it is telling you what each team is doing.

It turns out that a network property called Betweenness Centrality, which I won't define because they didn't give me a whiteboard. See, next time, have a whiteboard! Betweenness Centrality, if I calculated the betweenness centrality of each team and I look at which team is bigger, I can predict which team is going to win and I can predict which team is going to win better than I can from any of the other statistics that Fox Sport gave us when we were looking at the data.

These network measures are actually a measure of how effective the game is and how well you are doing in moving the ball from one side to the other.

Why am I standing here instead of sunning in the Bahamas and retiring on my millions that I made off betting on Australian Football?

First reason, I don't really believe in betting but the second reason, you have to collect all this data first. So, the betting agencies are collecting this, if you look at the football players, they all instrumented with the little things on their backs and someone's counting all the passes that they are making but they only gave me the data after the season and I was making these predictions after … well, this was a couple of seasons ago … it was after Fremantle lost again!

The point is that you can make these measures of these systems and predict the health and how these systems are performing. The trick is being able to do it fast enough to actually make money. That's a separate problem.

Let me finish and get back to the main applications that I was talking about at the beginning, mental health and maintenance and safety.

Within the Young Lives Matter project, we have started to look at trying to build models, mathematical models of two parts of the mental health process.

One is trying to make predictions of peoples state through time which is a data science problem I will describe briefly, but the other is actually looking at the health system. So, when someone presents to an emergency department, how are they treated? How are they interacting with the system? And, just like with the football passing through the football team, how is the information relevant to that patient moving about the hospital and the various other information systems to that person.

So, if someone presents to the hospital, they are in some sort of mental health crisis and the Registrar, the treating psychiatrist, they need access to his information. How do they get that and how is transmitted between people?

We did a study at one of the hospitals here, one of my post-docs in mathematics, I gave him a clipboard and sent him up to the hospital emergency department and he followed people around, with their permission. He followed people around as they were admitted to hospital, logging their interactions with the health system and what we found, I can't tell you the full story, but what we found was that certain individuals, with individual roles within the hospital had a key role in transmitting the information. So, without this person in the particular role in the hospital, the information between the emergency department and the patient and the health systems that store the information about the patient, that information did not get across and more to the point, if we took this person out of the system, the consultant psychiatrist had to do a lot more work. We presented these results to the hospital and they were looking at trying to figure out a way to optimise the system so your information works better.

The other side, the flip side, trying to predict what people are doing, that is more complicated.

I won't bore you with too much of the details there, but the simple thing about it is … are they any psychologists or psychiatrists in the room?

Okay.

So, psychologists and psychiatrists don't do very much. [laughing] What I mean by that is, if they are trying to assess people, they ask them various questions around, how are you feeling and how do you feel about this, this or this? And this is the data. It is a bunch of numbers between one and seven based on how happy someone is.

The trouble is if I asked Leo on a scale of one to seven and if I asked Thomas on a scale of one to seven, their idea of one to seven is completely different.

Individually, those numbers tell you nothing.

What we are doing is looking at how the individuals rating changes over time. It is not a radical idea but it is important to look at how and individual person changes their view of the world through time to make a prediction of how they feeling and how they are interacting.

Lastly, let me finish with the last application just so that our partners in the resource maintenance hub, a shout out to BHP, Alcoa and Roy Hill. There, I have done that! Our partners in the resource sector get their fair due and so we are interested in maintenance in the resource sector trying to predict when machines fail, but we are also interested in safety.

In safety systems, you have this management model to deal with. The industry process to make sure things go well and this a flowchart. It's a network. It's a bunch of friends connected by links and there

are points in that network which prevent bottlenecks. It is bottlenecks that constrict the flow of information, so we are trying to identify those bottlenecks and better optimise those networks.

In maintenance, we are looking at exactly the same model, using this wonderful piece of software called SAP. I think some of you have heard of it, maybe even used it. It is horrible. I hope no one from SAP is here. Anyway, we are looking at the various logs of jobs in the maintenance sector, in the maintenance part of the SAP for these industries and how different people are acting. So who is doing what work, when and where and how is that work distributed? Where are the bottlenecks? Where are the lynch points? And how does the affect the overall efficiency of the system in exactly the same way that we are looking at the efficiency of hospital systems or football teams?

I am going to stop there and open up for any questions.

[clapping]

Simon Handford

Thanks, Michael, really interesting.

The first thing I feel as though I should be doing is offering to buy Tom a drink. [laughing]

Professor Michael Small

You already did actually! [lauging]

Simon Handford

I am pleased to hear that!

Have we got any questions for Michael? We have got a few minutes, we have easily got fifteen or twenty minutes if there are some questions. We would love a bit of interaction from you guys in the audience.

I was curious whether any of these, mathematically driven immunisation strategies had taken place yet? Has that been deployed? With things like Sars and other outbreaks?

Professor Michael Small

Yes.

The most prevalent place where it has been deployed is actually in something called “ring immunisation” in diseases like Foot and Mouth in cattle. Of course, you are not asking the cow to nominate a friend but you are essentially putting a ring around the outbreaks in essentially the same way.

In human transmitted diseases, a lot of the effort goes into identifying the super spread as the hubs, the high connected points and trying to immunise them and this is one way to achieve that. Essentially, you can actually climb a hill. So, I start with someone who is only average, Thomas.

Simon Handford

It's going to be Thomas again, isn't it? [laughing]

Professor Michael Small

I asked Thomas to nominate one of his friends and on average, that is going to be someone more popular but there is another person out here who is also named Thomas who is a friend of Thomas, then this Thomas for one of his friends and we haven't got Thomas number three here, have we? T3 is not here, but one of his friends on average would be more popular! So you can climb up and find the hubs that way and this augments all the other sort of intelligence that you might use to try to identify the locations and individuals that are highly transmissible.

Of course, we all know healthcare workers are a key to this because they are dealing with a lot of people, a lot of sick people, things like flu.

Primary school teachers, they are the first, the frontline there. So, a lot of it is known already.

Simon Handford

Any questions from the audience? Anybody working in a sector that might benefit from some mathematical modelling? Any of you working in complicated areas with nodes and networks and that sort of thing?

We have a question here.

Audience

I am asking for a friend, are there any lotto numbers for today? [laughing]

Simon Handford

Any lotto numbers for today Michael? I think we have missed Thursday Powerball.

Professor Michael Small

So, I forgot to mention my other side project.

This is where you give me fifty dollars and I give you six numbers. It is six, right?

Predicting lotto is a bit pointless but things like financial markets you can apply these sorts of things to try and predict that. In all honesty, it is a lot of work because a lot of other people are trying to do

exactly the same thing and in the financial market, in particular, anything that works is immediately washed out by the market itself, so sorry.

Simon Handford

I was going to give you fifty bucks.

Professor Michael Small

I will still take the fifty bucks. [laughing]

Simon Handford

I thought you might.

Any other questions?

Professor Michael Small

So the question is about any ethical issues about who gets vaccinated.

My cop-out answer is, I am a mathematician. I am not dealing with ethics.

But that is a bit of a cop-out. In any area where you have data, there are ethical issues that come up. Not to take that example but to take a different example, a lot of data mining strategies for example, because the data that you are using to build your model is biased and maybe it consists mostly of men. Your model that you get out of it is biased and there has been really well-publicised examples of certain speech bots that certain tech companies try to produce the immediately turned out to be racist, misogynistic, all sorts of other bad things, so you do need to think about the ethics of it, with any research.

You are producing a new idea, and that hopefully is going to have an impact and if something has an impact then there is potential that it can be used in many different ways.

Specifically for immunisation, the biggest thing that I would like to talk about and I would take more than twenty minutes to do so, is the issue around the movement against vaccination. The anti-vaxer movement and this is interesting from my perspective because a lot of the … you have two different sorts of transmission of information here. You have transmission of the disease that people are trying to vaccinate against and by the way, vaccination is good, measles, mumps and rubella have nothing to do with autism.

People are vaccinating to prevent the spread of disease but information is being spread on social networks that go against that and a lot of it … so there are people … there are clear cases of certain pockets of resistance to vaccination in certain neighbourhoods but a lot of it is also spread over social networks because someone is posting this information on their Facebook page but I won't delve into that right now because I will go on forever about that.

Simon Handford

I reckon that could be a talk for this time next year.

I was going to ask what are you most proud about in terms of your research and your research team and the impact that you have already had. What you have learnt and what you have deployed?

Professor Michael Small

That's a good question.

Why didn't you give me these questions before, so I had time to think about them and give you a nice pithy answer?

Simon Handford

Sorry.

Professor Michael Small

Honestly, personally, it is actually something … it's not directly related to research but it is seeing my students go on to be a success in their own careers having trained PhD students, Honours students and in fact today we just had our Honours presentations in the Maths Department seeing young people develop those skills and go out into the real world. Even Thomas!

I won't take that one, I will leave that one out there Thomas.

In terms of research, we did some really nice work around Sars, showing that the way Sars is spread is actually … it can be described exactly the sort of social idea.

A lot of the epidemiological research is talking about how you have "super spreaders" you have special people who are extra contagious and they are the cause of the problem. But actually it is not the people who are extra contagious, it is just that some people have more contact than others and that variability of contact is enough to explain exactly the distribution of spread that we saw with Sars.

I hope when you ask me that question in five years' time I can say something about the "Young Lives Matter" project. Unfortunately, the results for that are at the early stage.

Simon Handford

Anymore from the back of the room.

Oh there is a question here, one from a little bit closer

Professor Michael Small

You said the back of the room, that's not the back of the room?

Audience

So, using the idea that all of us use social networks and have access to information, what about developing a model to predict what a successful student is? [laughing]

Professor Michael Small

Wow, that is a good question and have to be very careful in phrasing my answer to that as well.

I guess, yes, you could identify good behaviour, good practices, things that on average are indicators of success and you could identify things that are indicators of failure … I guess failure is the wrong word, I should be saying, "not a success." Whatever the correct way of saying that without saying failure, is?

I am really … I really don't know … I mean my youngest eight and my eldest is sixteen and he is going into [inaudible 45:17] and all that, and he is listening to this absolute God awful, crap music all the time, driving me insane! But he reckons it helps him study. Maybe it does. It doesn't help me do anything but individuals are different.

The mental health project, yes, we are looking at individuals, we are trying to look at indicators just like this so there is a reason to think you could do the same but it is complicated.

I think a lot of it would actually be directed at the system.

Turn your question around and say what does the system have to do to make students successful.

That I don't have any answers for yet, either. [laughing]

Audience

[inaudible 46:06]

Professor Michael Small

By predictors, do you mean the data that goes into the model or the thing that we are trying to get out?

The things that go into the model at the moment, we are constrained by what's available.

There are lots of things that we might like to measure for ethical reasons or for practicality reasons, they are very difficult to measure and the other thing that happens here is that you are trying to predict something which, although a big problem, is still, fortunately, a fairly rare event. So, you can't really just go after targets and look at them, you have to look at every one. So, you are limited by practicality, what you can do to large populations. So, in the study that I was talking about, we

were working at an in-patient psychiatric hospital in Perth and they were collecting these psychiatric surveys on an almost daily basis so that was the main data coming into it.

There is a lot of standard population type statistics that you can get as well, the usual things, socio-economic conditions, home conditions, drugs, alcohol, all those things. They are predictors but they are not very good predictors on an individual level. They predict populations. What we are going after is ways of trying to assess individuals v populations.

At the moment it is Likert-Scale tests where on a scale of one to seven, how do you feel about whatever, that is most of your data.

We are hoping for more and we are hoping to be able to identify things that we can try to measure in the future but it is a complicated process.

I poo-pooed the education but mental health is far more complicated.

Simon Handford

Any final questions for Michael before we wrap things up?

It has been fascinating for me to hear about all the mathematics and everything from disease control and vaccinations to oil and gas and mental health, it underpins so much of what we do.

Professor Michael Small

I didn't connect mental health and oil and gas. [laughing]

Simon Handford

I am glad you clarified that!

Can we just thank Michael one more time?

[clapping]

Professor Michael Small

Thank you for listening.

Simon Handford

I think you have made a few friends yourself here tonight, so thanks very much.

[clapping]

We have another talk coming along soon,

If you have registered for that one, you are welcome to stay, if you haven't we might be able to squeeze you in but it might be good to just go downstairs and stretch your legs and then we might just check the numbers as the people who have got tickets for the second session come along and register.

Thanks so much for coming.

This has been Raising the Bar and UWA.

Thanks a lot!