Quantitative power-mapping: a proof of concept in Boston

Rashid Amerzaine
Bernard Flores
September 6, 2023

1 The scenario.

We ask the reader to imagine that, in spite of the many obstacles in the way[1], they are a member of a healthy, consensus-based community organizing group using direct action and mutual aid in the struggle against settler colonialism, imperialism, and militarism. It is well known to the membership that the forces they struggle against are maintained by vast networks of political collaborations, business relationships, and capital investments.

The group has limited capacity and needs to be strategic about how it designs a campaign for accomplishing its goals. For example, perhaps after hearing that many members are being priced out of their neighborhoods as rents soar higher, the group sets its sights on housing justice. To set strategy, the group engages in the standard organizing practice of *power mapping*: delineating the key players who have influence and the capacity to either maintain the status quo or shift it, and where they stand in terms of being either in our out of alignment with the group's organizing goals.

During the exercise, many potential targets for a campaign emerge:

- the developers who buy up land to build fancy condos and cut further into the already dwindling supply of affordable options;
- the landlords who hike up rents at any conceivable opportunity;
- the business owners who spur gentrification by catering to yuppies, calling the cops on longtime community members, and colonizing community spaces with irrelevant, expensive, and exclusionary establishments, eliminating possibilities for non-commercial existence in public;
- the banks that help to maintain de facto apartheid through biased credit risk models and predatory lending;
- the realtors, property managers, and other parasitic intermediaries who benefit from high rents and frenzied turnover, and run interference for landlords;
- the politicians who write laws protecting the landlord and the bank at the expense of the renter;

[1] At least in the US, resourcing and building an organizing group that is genuinely led by a strong base of community members and that is dedicated to anti-imperialism, racial/gender/sexual justice etc. is extraordinarily difficult. There are many hurdles to clear all while having to avoid elite capture and other pitfalls of the non-profit industrial complex. Having said that, we firmly believe in the power of grassroots community organizing. Fear of its potential is precisely why the system places such large obstacles in its way.

- the bailiffs and police upon whose threat of physical force all of the above rests.

These players exert many forces, occasionally opposing each other, often coordinated, sometimes intentionally, sometimes not. But the sum of those forces is: human beings – members of the community organizing group – being forced, or at risk of being forced, from their homes and communities.

So, *how does the group choose where to focus?* Should they try to shut off one of the motors generating those forces, by confronting one of the above players? If so, which one? Or should they clog the gears of the machine, by disrupting the connection between some pair of the players? If so, which pair? Our goal in this piece is to explore these questions with the help of some basic mathematical tools of network analysis: *centrality measures*.

We want to be careful at this juncture to avoid a pitfall common in the "mathematics applied to social systems" literature: we are not interested in framing this question as a purely mathematical one. Mathematical models are always predicated on a choice of features of the situation being modeled, namely which ones will be abstracted in the modeling and which ones will be ignored. So, any mathematical analysis of the above question may ignore exactly the set of contextual, historical, or social factors that make some target more attractive in actual practice than the one recommended by the model.

Regardless of what a mathematical analysis reveals as the "optimal" avenue for directing scarce organizing energy and resources, perhaps there happens to be[2]

a huge amount of energy and community support for some other campaign. In this case, the organization might opt to move in this other direction, because choosing a campaign for which there exists significant energy is a strategic way to minimize burn-out and secure the necessary capacity to achieve the campaign's goals.

What we hope to offer is proof of concept for what might be one of many tools for supporting complicated collective decisions in the context of a movement. We don't want to "mathematize" the entire problem; such an exercise is as uninteresting and useless to us as it would be to organizers. We want to stimulate thinking about the possible role of mathematical techniques, deployed in narrowly-specified ways, in organizing strategy and tactics. And we want to discuss this in the context of plausible examples.

2 Who we are.

We are writing this as members of the *Just Mathematics Collective (JMC)*, and one of us – Rashid – previously co-authored a piece[3] with the same pseudonym. In that piece, they made clear who they were, and this description applies here as well: we are mathematicians with permanent academic jobs in the US. We don't claim access to knowledge of experience outside of these positions and we speak only for ourselves. Our choice to write anonymously comes out of a general JMC ethos of collectivity; this may have been written by any pair of people in our predominately anonymous membership. Having said that, while collective members have had the opportunity to read and comment on this piece before it was released, it does not follow that if someone is a member, they agree lock-step with all claims and contentions.

Both of us have experience in movement spaces. We have both belonged to organizations like the one described above: community-led, consensus-based, and focused on fighting

[2]"happens to be" does not mean "randomly and for no reason", it just means "for no reason that's not been flattened in the process of mapping reality to the model"

[3]*No cops in the lecture hall: cheating and what (not) to do about it*, `https://www.justmathematicscollective.net/cheating.pdf`.

against systemic injustice. Having said that, **we are not "experts" of community organizing**, a concept we'd reject. Before we expound on this, we want to make clear that of course there are people with *vast* amounts of experience and knowledge about how to build power in their own communities, people who've forgotten more than either of us will ever know. We both know such people and are constantly learning from them and do our best to follow their leads.

Having said that, we believe strongly that the most powerful organizing results when individuals feel empowered to share freely, exchange perspectives, and communicate directly, all while of course acknowledging the differences in experience that surely exists across a diverse group of people. Power comes from each person feeling clear on how their own interests align with the goals of the collective and feeling confident in their capacity to articulate those alignments. Each person, having connected their own interests to the collective goals, and having had a real opportunity to express their perspective and (actually) influence decision-making, then commits to collectively-decided courses of action.

So to summarize: we are mathematicians who believe in the power of community organizing, writing for a predominantly mathematical audience; on the other hand, we make efforts to write in a way that is as accessible as possible to people without any formal mathematics training. We have zero interest in telling organizers how to "optimize" in their movement building. We hope that this piece encourages mathematicians who are involved in their own organizing spaces to think about

- how their specific knowledge and training might be useful during campaign planning, and
- how to present abstract mathematical tools to students/friends/comrades in ways that demonstrate their capacity to build collective power.

3 Our politics.

It feels important to both of us that we are forthright about our political beliefs and commitments. To be clear, the analyses we outline below and the tools of quantitative power mapping apply to any number of movements. But – perhaps especially because Palestinian liberation, the unavoidable subject of this essay's main example, is such a touchy and censorious subject in academic spaces – explicitness about our stance on the matter is imperative:

We believe in and we hope for a free Palestine from the Jordan River to the Mediterranean Sea. This means that we believe in a future for the region where all people are treated equally (regardless of religious/ethnic/racial identity) under the law, have access to self-determination and freedom of movement, and have the right to return to the land from which they and their families were forcibly removed.

The Nakba – often marked by the ethnic cleansing of hundreds of thousands of indigenous Palestinians in 1948 – arguably began decades earlier, and has never ended. To this day, Palestinians are subject to brutal occupation, military blockades, and statelessness. We wholeheartedly support the movement for Boycott, Divestment, and Sanctions (BDS) spearheaded by Palestinian civil society. Beyond this, we affirm the right of armed resistance for every occupied people. In spite of this right being enshrined by international law, many would deny it to Palestinians, instead opting to demonize any Palestinian who picks up arms to defend themselves and their families from an invading colonial force as a "terrorist"[4].

[4]In our circles, comparatively very few people object to the affirmation of this ethical and legal principle in

We refuse to avoid phrases that bad faith reactionaries, such as the Anti-Defamation League[5] (we strongly encourage readers to peruse the last footnote as an illustration of this organization's untrustworthiness and reactionary agenda) have tried repeatedly to criminalize and/or paint as hateful or antisemitic. We are grounded in our own understanding of Judaism and that of many of our comrades, and we know with our full hearts that anti-Zionism is absolutely *not* antisemitic. To the contrary, we relate to our anti-Zionism through a longstanding tradition *within* Judaism that honors justice, liberation, and anti-authoritarianism. The fact that many Jewish people who speak up for Palestinians end up as targets of reactionary pro-Israel campaigns suggests that propagandists for the Israeli state want nothing more than to sever people from their own histories and from these culturally legitimate and important ways of being Jewish.

We refuse also to avoid these phrases (for example "From the river to the sea, Palestine will be free", or from referring to Israel as an apartheid state, or to the ongoing Nakba as a genocide, or from stating our affirmation of Palestinians' right to armed resistance) out of respect for the clarity and forthrightness of those who have said such things before us. When we first began planning this piece, we entertained the idea of pulling back a little and being a bit more coy or "strategic", but we came to the conclusion that this is ultimately a losing tactic which serves to further alienate those honest enough to have made their commitments clear in the past.

As we began to write this, a refugee camp in Jenin was under siege by Israeli forces. People were being forced at gunpoint to leave their homes– thereby becoming refugees twice over. Israeli soldiers were intentionally destroying civilian infrastructure and blocking access to essential medical resources [6], and Palestinians were being targeted brutally and indiscriminately. The urgency of the situation on the ground is not compatible with further indecision on our part about what linguistic choices are rhetorically optimal.

So in short: we will not be drawn in by reactionary hasbarists, and we feel that our convictions are adequately described by the above and by the statement made by the Just

---

the case of the Ukrainian people's self-defense against Russian imperialist aggression, while many of the same people would take all characterizations of all instances of Palestinian resistance as "terrorism" at face value. Distinguishing freedom fighters from "terrorists" is a pursuit obviously rife with special pleading and double standards.

[5]The Mapping Project – the group whose work inspired this piece – wrote a carefully sourced article (`https://tinyurl.com/bdecah7b`) about the history of the ADL and its pattern of sidling up next to police and state surveillance agencies, throwing even many Jewish people under the bus when it serves their true interest of promoting Zionism and the Israeli government. We briefly mention some highlights: (i) the precursor of the ADL sent millions to the Jewish National Fund in the '30's and '40's to support land theft and dispossession of Palestinian homes. (ii) In the wake of the Bolshevik revolution, the ADL launched a media campaign meant to imply that all American Jews were anti-communists. (iii) It stood by quietly during Japanese internment in the '40's. (iv) It spied on "dissident" Jews during McCarthyism in the '50's and handed over its "findings" to the FBI, and even went so far as to warn Jewish orgs not to support the Rosenbergs in any capacity when the US government accused them of espionage. (v) In response to a 1967 article written by the Student non-Violent Coordinating Committee (SNCC) correctly characterizing Zionism as imperialism and colonialism, the ADL accused SNCC of being a "Negro extremist" organization and "no longer a responsible civil rights group". (vi) In 2016, the ADL attempted to tar the Movement for Black Lives by denouncing it as anti-Semitic. (vii) For decades, the ADL has worked hand and hand with US and Israeli governments to spy on Arab students, communities, and congregations. (viii) For over 70 years, the ADL has been involved in promoting and training police. Post-9/11, they've ramped up efforts to fund and promote police exchange programs, whereby American cops are sent on all-expense paid trips to Israel to learn how to stifle dissent, revolution, unrest, etc. most efficiently. (ix) The ADL promoted US propaganda in Latin America, eg. parotting US concerns in Chile about Salvadore Allende's socialist government. In doing so, it characterized all Jewish people in Latin America as capitalists (a deeply anti-semitic claim) as a means of suggesting that a turn towards socialism would be dangerous for all Chilean Jews.

[6]See the following article from *Doctors without Borders*: `https://www.doctorswithoutborders.org/latest/msf-condemns-denial-medical-access-during-israeli-military-raid-jenin`

Mathematics Collective when it launched its own BDS campaign within mathematics[7].

4 Background on the Mapping Project.

This piece was inspired by the so-called *Mapping Project*, a collective of organizers and activists who compiled an extensive dataset and released it publicly in the spring of 2022. The Mapping Project explain themselves and their motivations perfectly well on their website:

> *We are a multi-generational collective of activists and organizers on the land of the Massachusett, Pawtucket, Naumkeag, and other tribal nations (Boston, Cambridge, and surrounding areas) who wanted to develop a deeper understanding of local institutional support for the colonization of Palestine and harms that we see as linked, such as policing, US imperialism, and displacement/ethnic cleansing. Our work is grounded in the realization that oppressors share tactics and institutions – and that our liberation struggles are connected. We wanted to visualize these connections in order to see where our struggles intersect and to strategically grow our local organizing capacities. Our interactive map illustrates some ways in which institutional support for the colonization of Palestine is structurally tied to policing and systemic white supremacy here where we live, and to US imperialist projects in other countries. Our map also shows the connections between harms such as privatization and medical apartheid, which are often facilitated by universities and their corporate partners. Since local universities engage in these multiple forms of oppression and produce much of the ruling class, and because they are major land holders in our area, we've emphasized the university as a central nexus that ties together many of the harms traced on the map...*
>
> *...We acknowledge that our map is not a complete representation of local institutions responsible for the colonization of Palestine or other harms such as policing, US imperialism, and displacement. We also recognize that the struggles of local Indigenous nations against US colonization are under-represented on our map. We would be grateful for suggestions and knowledge shared with us by those who engage with our map, and hope it can continue to grow and improve through your contributions.*

As mentioned above, the project consists of an interactive map involving almost 500 different entities operating in or near Boston, Massachusetts. The user can hover over any of these nodes and learn why that particular institution was included in the network and how it furthers the forces of Zionism, imperialism, policing, gentrification, or medical apartheid. The user can also toggle between different types of relationships shared by these organizations, like "collaboration/association" or "financial support". Our proof-of-concept analysis of Mapping Project data focuses on the financial relationships.

4.1 Accusations of antisemitism.

Because we don't believe the slanderous allegations directed at the Mapping Project, we don't want to spend too much time on them. But to provide context and summarize the controversy that brewed around the project, it is necessary to outline the accusations.

When the project went public, many reactionary voices[8] claiming to represent the "Jewish

[7] See `https://www.justmathematicscollective.net/USACBI_statement.html`.

[8] See for example the following which argues that the "BDS Mapping Project" (conflating two totally distinct campaigns) is dangerous for American Jews, and everyone, for multiple reasons including that it represents an "ideological attack on Western values", characteristic of the "Neo-Marxist ideas that surge through this movement": `https://www.ajc.org/news/5-reasons-why-the-bds-mapping-project-is-dangerous-for-jews-and-everyone-else`.

community"[9] sounded an alarm: *this interactive website lists the names and addresses of Jewish people and Jewish organizations and calls for the "dismantling" of the network it maps.*

It's important to point out that many of the organizations upon which the project focuses do not identify as Jewish in character. Indeed, as we will see below, some of the most centralized nodes in the network– the nodes which arguably contribute to the forces mentioned above more than any other– are major universities like Harvard. Nevertheless, it is true that there are certainly a number of Jewish groups and organizations included in the network, and this should not be surprising. While of course there are many Jewish organizations and congregations who do not make a point of supporting (in rhetoric or in material resources) the Israeli state and also those who make a point of *not* supporting it, there are also many that do.

It is **crucial** to understand that among the Mapping Project's criteria for inclusion of an organization in the network was that organization's **explicit support for Zionism**. Their criteria did **not** allow for the inclusion of an organization on the basis of its relationship to Judaism or Jewish people. An observer committed to engaging honestly with actual reality will not find this distinction difficult. However, people with a vested interest in advancing the slanderous, and in fact antisemitic, line that to be Jewish is to be Zionist will, of course, want to elide the objectively obvious distinction between the Mapping Project's actual, non-antisemitic inclusion criteria and an imaginary, antisemitic inclusion criterion articulated by an equally imaginary straw-man. This is apparently what various organizations and commentators did, and it was, unfortunately, effective.

The next, much more subtle, comment emphasizes how confusing the discourse around this issue becomes when it gets refracted through political organizations and corporate or state media:

We acknowledge that the terrible history and inexcusable persistence of antisemitism in Europe and in the United States will naturally inform the perception of a project like this by some Jewish people[10]. The inter-generational trauma of pogroms, scapegoating by power-hungry Christian demagogues, the history of institutional discrimination, old and recent, and, of course, the Holocaust, shapes a wholly reasonable and justified cultural wariness about being singled out or blamed for larger societal ills. We understand these fears, *and at the same time*, our empathy with Jewish people who may feel this fear coexists with our firm belief that the Mapping Project does not, in fact, aim to single out Jewish people for scapegoating or blame and is not, in fact, antisemitic. Further, conflating the Mapping Project with an antisemitic undertaking – though factually wrong– needn't have happened maliciously in every case, especially considering the significant efforts by (Israeli, American, and European) state actors to cynically play on valid cultural fears.

Moreover, it is our contention that the Israeli state and organizations like the ADL very frequently exploit these traumas to manufacture consent for their reactionary agendas: it is in the best interest of the Israeli State for Jewish people to be afraid of any anti-Zionist project[11]

[9]The quotes here are to emphasize that Jewishness across the world is not monolithic. There are a myriad of distinct Jewish communities that differ from one another in almost any dimension of daily life, beliefs, and cultural practice. Nevertheless, there is no shortage of groups– for example the ADL– claiming to represent all Jewish people, and since the ADL is staunchly Zionist, in attempting to speak for all Jews it furthers the harmful trope that to be Jewish is necessarily to be Zionist.

[10]And by people, Jewish or not, who are aware of the need for vigilance in service of eliminating antisemitism.

[11]We invite the reader to explore *Decolonize Palestine*, a primer on settler-colonialism in Palestine which covers, among many other things, some of the propaganda tools used heavily by Israeli state actors: `https://decolonizepalestine.com/`.

In any case, the Mapping Project made national news and the accusations of targeting Jewish people swirled around it for months. Other organizing groups that claim to share many of the same goals as the Mapping Project denounced it, we speculate, to avoid being tarred with the same brush. We characterize the affair as an instance of censorship both of the Mapping Project specifically, and as part of the ongoing censorious maintenance of a chilling effect on discussion of solidarity with Palestine. As a consequence, the project's utility – both as part of the movement for supporting the Palestinian liberation struggle, and as a tool for organizing around matters at some remove from Palestine – hasn't been properly explored (at least in our humble opinion). Our motivation for this piece was to make a small contribution in this direction.

5 The (very) basics of centrality measures.

This section is written under the assumption of a reader familiar with certain mathematical notions, but will aim to be as accessible as possible.

We will use blue text when writing in a mathematically formal way– anyone who does not feel practiced enough with this sort of writing to get through the blue bits is encouraged to skip them: we have tried to structure this so that someone skipping the blue bits will, hopefully, still get a reasonable picture of centrality measures, adequate for this discussion, by reading the rest.

A *network* is a collection of *nodes* – you can safely think of nodes as dots on a page that might represent specific things; in our case, you can think of a node as a dot on a page representing a Boston-area institution – some pairs of which are connected by *edges*. You can think of an edge as a line segment on the page, connecting two of the dots.[12]

The edges may be decorated with arrows pointing away from one node and towards the other; in this case, the network is said to be *directed*. For instance, if nodes are Boston-area institutions, we might put an edge between two of them to indicate the presence of a financial relationship. We might make it a directed edge to indicate the direction in which money flowed.

This very simple and versatile structure is used to model many phenomena; well-known examples include

- social networks (nodes represent people, and an edge between two nodes represent friendship or some other sort of relationship)[13];
- biological systems, e.g. the brain (nodes represent neurons, and an edge represents a synapse between a pair of neurons);
- traffic patterns (nodes represent landmarks in a city or state; edges represent major arteries or roadways connecting nearby landmarks. In the event that these roads are one-way, the edges may be directed).

Network analysis applies in an innumerable variety of other situations.

We will denote a network by $G = (V, E)$, where $V = \{v_1, ..., v_n\}$ is the set of nodes and $E \subset V \times V$ the (directed) edges. For undirected graphs, the edges are a subset of the quotient of $V \times V$ obtained by identifying all pairs (v_1, v_2), (v_2, v_1).

A *centrality measure* is a way of quantifying how centrally positioned each node is in a given network.

[12] It doesn't matter how long, straight, wiggly, etc. the segment is. It's just a representation of the presence of a specific relationship between those two nodes.

[13] We direct the reader to the following classic example in sociology wherein network analysis is used to study the rise of the Medici family: `https://home.uchicago.edu/~jpadgett/papers/published/robust.pdf`

Formally speaking, one can think of a centrality measure as a function $F : V \to \mathbb{R}_{\geq 0}$ which assigns to each vertex a non-negative real number; higher numbers correspond to nodes that are located more centrally in the network.

Here are just a few centrality measures; we refer to Section 5.1 for concrete examples showing how each of these is computed and how they compare, in the artificial example of a very small network with eight nodes.

1. **Degree.** Assign to each node the number of other nodes that are connected to it by an edge. If the network is directed, we can also speak of *in-degree* (the number of nodes pointing towards a given node) and *out-degree* (the number of nodes towards which a given node points).
2. **Closeness.** We can define the distance between any two nodes in a network to be the smallest number of edges one must traverse to go from one to the other along an unbroken path in the network, i.e. start at the first node, move to one of the *adjacent* nodes (nodes joined to the first one by an edge), then move to a node adjacent to that one, and continue in this way until you get to the last node. Count how many edges you "traveled through". Imagine all the possible ways of doing this, and keep track of all those counts. The smallest one is the closeness.

   In the event that the network is directed, to go from one node to another we must travel along each edge only in the specified direction; it's possible in this setting for the distance from some node A to another node B to be different from the distance from B to A. This reflects the reality of many situations in which the appropriate notion of distance is not symmetric– consider for instance a city with many one way streets, or even simpler, that it's generally a lot harder to go up stairs than down them.

   The *closeness centrality* of a node A in a network is the reciprocal of the sum of distances from A to every other node. If A is very close to most other nodes and B is relatively far away from most nodes, the closeness centrality of A will be larger than that of B because, to calculate the closeness centrality of A, we divide 1 by a smaller number than we would divide by when computing the closeness centrality of B. Formally speaking,

$$C(A) = \frac{1}{\sum_B d(A,B)},$$

   where $C(A)$ denotes the closeness centrality of the node A[14].

   In some literature, the closeness centrality is defined a bit differently to be 1 over the *average* distance of a given node to every other node. To compute the average distance from node A to all other nodes, we simply add up all distances between A and every other node, and then divide by the number of nodes. It follows that this alternative definition of closeness centrality is given by the formula

$$C(A) = \frac{\#(\text{nodes})}{\sum_B d(A,B)}$$

   This alternative formulation is related to the first definition by just dividing by the total number of nodes; Mathematica uses this second formulation, so we will see it in action in the later section on findings.
3. **Betweenness.** Those familiar with the dynamics of the American airline industry are acutely aware of "hubs" – an airport chosen by a particular airline to be an unofficial

[14] First of all, note that $d(A,B)$ can be different from $d(B,A)$ because paths can be directed. Secondly, in the event that there are multiple connected (directed) components, our convention is to compute closeness within each component separately.

headquarters, to and from which a disproportionately large number of flights arrive and depart. This allows the airline to save time (and ultimately, money) organizing the logistics of its itineraries. Of course, the time saved by the airline is often time stolen from the passenger, who–if for instance they are flying Delta – may find themselves stopping for a long layover in Atlanta on their way to Alaska from New York. This is perhaps one of many losses the passenger experiences as a result of the relative power imbalance between an industry selling the increasingly inelastic product of long-distance travel. (Facilitated by globalization and often built on a foundation of settler colonialism[15]. But we digress.)

In the network where nodes correspond to airports at which, for example Delta operates, and edges correspond to Delta flight paths, the Atlanta airport would have a high *betweenness score*. Betweenness intuitively measures the number of efficient[16] paths in the network that pass through a given node.

Formally, the betweenness centrality $\mathcal{B}(A)$ of a node A is defined to be the sum–taken over all other pairs of nodes X, Y– of the proportion of shortest paths from X to Y that pass through A:

$$\mathcal{B}(A) = \sum_{X \neq Y \neq A} \frac{S_A(X,Y)}{S(X,Y)},$$

where $S(X,Y)$ denotes the number of shortest paths from X to Y, and where $S_A(X,Y)$ is the number of such paths passing through A[17]

4. **PageRank.** One can imagine a network in which nodes correspond to webpages on the internet, and there is a directed edge from site A to site B when A links to B, i.e. when a user can visit B by clicking on a link from site A. Search engines like Google use an algorithm to help determine the order with which to show search results to a user.

   For example, if a user searches for "climate change", ideally the first hit Google returns will be a highly reliable and well-evidenced source, as opposed to, for instance, the blog of a climate change denier. To solve this problem[18], the search engine treats a webpage as more reliable when many *other* webpages link to it. For example, one expects many more websites dealing with the topic of the climate crisis to link to scientifically-informed public bodies (e.g. the US EPA, or the Intergovernmental Panel on Climate Change), or perhaps to community-vetted non-commercial informational resources (like Wikipedia), than link to, say, a wingnut's Wordpress.[19]

   The very basic idea behind this algorithm is to assign an *importance score* to each website that correlates to the number of other sites that link to it; the higher a website's importance score, the closer to the top it will appear in a search.

   But what if our Wordpress wingnut has lots of money to buy hundreds of domain names and have all of those sites link to their blog, thereby artificially inflating the importance? In one sense, the answer to this question is "nothing", in that importance

[15]For example, the immense size of the United States and the fact that it is fairly common for someone to be hired far away from their family encourages air travel.

[16]By "efficient", we mean *with respect to* the network. So for example, passing through Atlanta on the way to Alaska from New York is certainly inefficient from an objective standpoint, if the measure is actual distance traveled, or fuel spent. But from the point of view of the network designed by Delta to maximize revenue and minimize logistical costs, the path through Delta flights that takes a passenger to Atlanta is efficient.

[17]A point of caution: if the graph is directed, note that $S_A(X,Y)$ can be different from $S_A(Y,X)$, and so we must consider both (X,Y) and (Y,X) as distinct pairs arising in the sum.

[18]The algorithms actually used in practice by commercial search engines are much more sophisticated than what we will describe here, but this exposition captures the basic idea.

[19]This is why propaganda that promotes climate change denial/dissembling through a massive institutionally supported online presence is dangerous to a degree that individual reactionary blowhards, acting alone, are not.

inflation is a major industry – search engine optimisation – and the rich have a number of ways at their disposal for manipulating internet searches to their advantage. On the other hand, this example suggests that importance should not be based solely on the number of sites that link to a given site[20], but instead, a site should be given more credit when websites *which are themselves important* link to it. This makes the Wealthy Wordpress Wingnut's job at least a little harder: if they want their army of shell sites to make an impact, they'll need to make some effort to increase the importance of each of *those* as well.

This makes the definition of importance sound circular: how do we define the importance of a site if it depends on the importance scores of many other sites? The mathematical approach to resolving this issue is a system of equations: one creates a variable for each website; let us denote the importance score of site A by X_A. We can then use a simple equation in these variables to capture the intuition that X_A should depend on

(a) the sites that link to A, and
(b) the importance scores of all of those sites.

So, if $B_1, ..., B_n$ are the sites which link to A and if $X_{B_1}, ..., X_{B_n}$ denote the importance scores of each of these sites, we can then set

$$X_A = \frac{X_{B_1}}{\text{outdeg}(B_1)} + ... + \frac{X_{B_n}}{\text{outdeg}(B_n)},$$

where outdeg denotes the out degree. The idea, in words, is that the importance score of A should be large when very important websites link to it. However, if a website with a large important score links to a huge number of other websites, A being only one of them, then the contribution it makes to the importance score of A has to be divided among these very many sites.

Doing this for each site gives us a collection of equations which we can then try to solve using basic tools of algebra[21]. For those familiar with linear algebra, this is equivalent to solving an eigenvector problem for the matrix of coefficients for the system of equations described above. We also rescale each column of the matrix of coefficients so that it sums to 1. This guarantees that the eigenvector we are searching for– one contained in the 1-eigenspace– will be Perron-Frobenius and will therefore be unique up to scaling.

Loosely speaking, a site has high importance when the chances that a user ends up visiting it are high, if they plan to bounce around from site to site by following links; at each site, they choose randomly among the links there to visit next. We can think of this as a kind of centrality measure and apply it to *any* network, not just web pages on the internet.

There are many more notions of centrality to explore[22]; the above is really just an arbitrary sample of the tools one can apply to study the positionality of nodes in a large network.

[20]The importance score is not simply calculated by counting the number of sites that link to a given site; that count is just the in-degree mentioned above.

[21]Solving systems of equations of this form is the bread and butter of *linear algebra*.

[22]As pointed out to us by a colleague who read an earlier draft and who knows much more about network analysis than either of us: one centrality measure that is currently gaining popularity and that we do not discuss is so-called "SpringRank" (see here: `https://www.science.org/doi/10.1126/sciadv.aar8260`. In this model, a directed edge from A to B is interpreted as an endorsement of A from B. Edges are treated like springs that are stretched in proportion to the difference in ranking between A and B. One solves for the ranking by minimizing the total energy of the system.

However, there are two crucial points to keep in mind about the variety of different centrality measures. We'll first state these two caveats, and then we'll discuss some examples, both to illustrate how some of the centrality measures work and also to demonstrate concretely what's meant by the two caveats. First the caveats:

- **(Conceptual independence of different centrality measures.)** The first caveat is that the different measures are expressed in different units and measure different things, so they can't be directly numerically compared. This is analogous to the fact that, say, several different numerical measures exist to summarize the current weather in a given location, but it's not generally meaningful to say "The temperature in my location is 14 degrees Celsius and the wind speed in your location is 12 kilometers per hour; you are having nicer weather than I am." Indeed, there is no single numerical scale of "nice weather" on which wind speed (in whatever units) and temperature (in whatever units) are directly comparable. This would be true *even if* "nice weather" were not a subjective notion. The measures "temperature" and "wind speed" might both be indicators of some more holistic notion of "weather niceness" but they aren't perfect proxies for it and aren't directly comparable.
- **(Verifiable independence.)** In many cases, the different centrality measures aren't just conceptually or formally unrelated, they are demonstrably independent of one another in the sense that one can make a real network where node A has, say, a higher betweenness score than node B, but B has a higher closeness score than A, while C has a higher PageRank importance score than B, etc.
  This is analogous to the simple observation that, when we look at a weather forecast showing tomorrow's temperature and wind speed in specific cities, it's likely that there will be concrete examples where the weather is "nicer" in City A than in City B, as determined by temperature, but "nicer" in City B than in City A as measured by wind speed.

These imply:

We need to be ***extremely careful*** *when we say that "A has a higher centrality score than B means that A is more central than B, or more important to the overall health of the network than B", or anything like these statements. Clearly, there are different and incompatible ways to quantify centrality, and there is no obvious way to prioritize these notions.*

One concrete form that being careful should take is: there needs to be a theoretical, conceptual, narrative connection between what the network represents or models, and what centrality measure we choose to prioritise. The different centrality measures capture different heuristic features of a network that are all defensible as notions of centrality or importance, but computing these measures has to be done in the context of an extra-mathematical understanding of the situation, so that we choose a centrality measure that's actually relevant to the real situation being modelled.

5.1 Centrality, and the two caveats, in a small example

Let's digress and discuss the caveats in the context of a concrete example. This section is mostly in blue text, but the reader comfortable accepting the two caveats above can skip this section without missing much.

Consider the network shown in Figure 5.1. There are eight nodes, which we've labelled $A, Y, X_1, X_2, X_3, Z_1, Z_2, Z_3$. The edges are directed, which is indicated by the arrows in the figure.

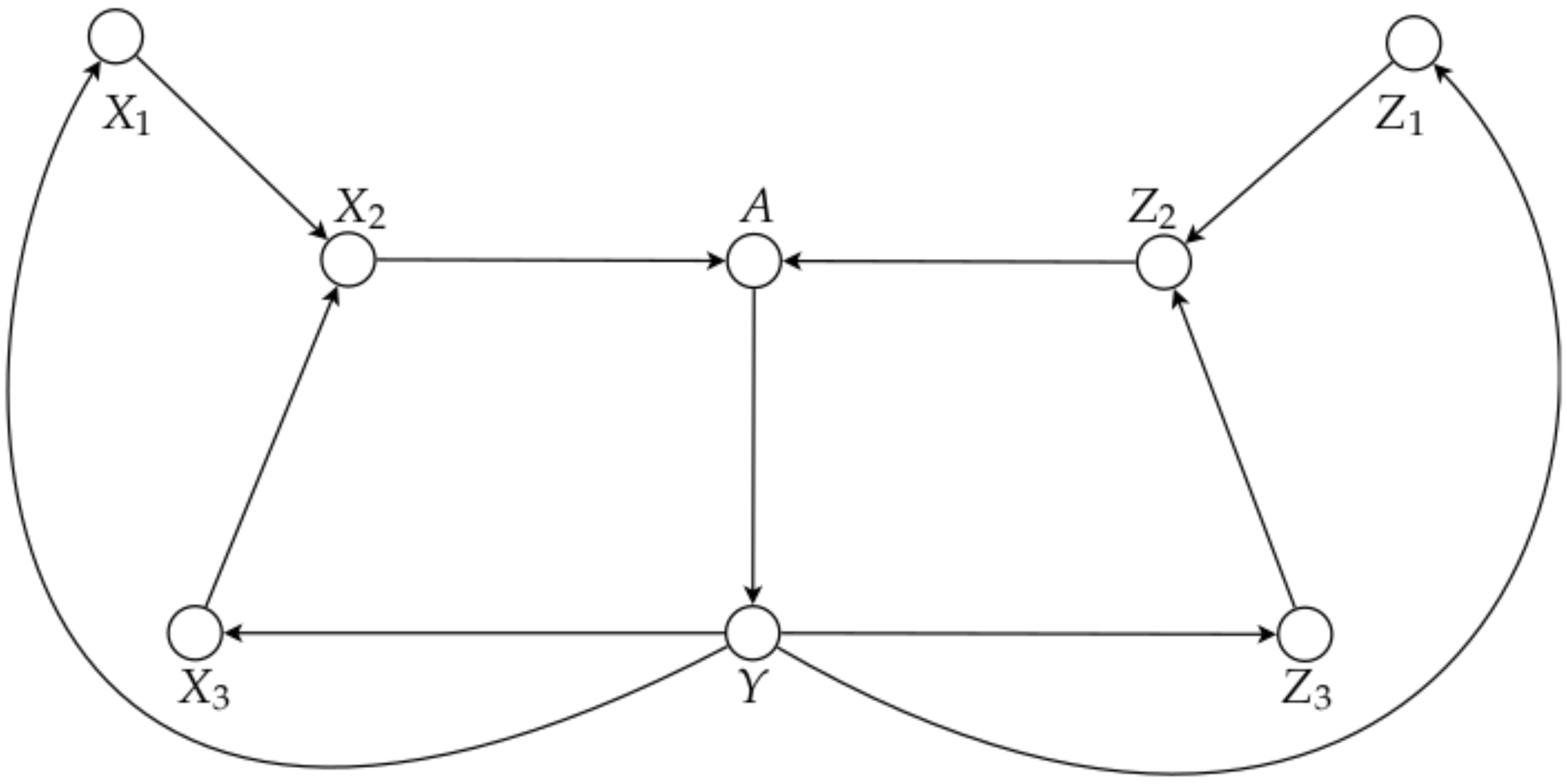

Figure 1: The (directed) network in Section 5.1.

We'd encourage the reader to have a look at the picture and try to use the definitions given above to calculate the in-degree, out-degree, closeness, betweennness, and PageRank for some nodes in the graph.

For instance, just by looking at the network and counting the arrows pointing into each node, and the arrows pointing out, we can see that, for instance, $\text{indeg}(A) = \text{indeg}(X_2) = 2$ while $\text{indeg}(Y) = \text{indeg}(Z_1) = 1$. The in-degree is largest at the nodes where it is 2. On the other hand, $\text{outdeg}(Y) = 4$, and out-degree is maximised at Y.

To compute the closeness is a little more complicated. Remember that, if P, Q are nodes, then $d(P, Q)$ is the length of any shortest path from P to Q, where we only allow ourselves the use of *directed* paths – we have to traverse each edge in the path in the direction of its arrow, otherwise that path doesn't count. So, for example, $d(A, Y) = 1$, but $d(Y, A) = 3$, since, to get from Y to A, the (non-unique) shortest path is Y, X_3, X_2, A, which uses three edges.

To actually compute the closeness centrality $C(Y)$ at Y, we have to consider all the other nodes, find the distance, and use the definition. So: $d(Y, A) = 3, d(Y, X_1) = d(Y, X_3) = d(Y, Z_1) = d(Y, Z_3) = 1$, and $d(Y, X_2) = d(Y, Z_2) = 2$. So $\sum_{P \neq Y} d(Y, P) = 11$, and we get $C(Y) = \frac{1}{11}$.

Repeating the computation for the other nodes, one can check that, for instance, $C(A) = \frac{1}{15}$ but $C(X_1) = \frac{1}{23}$, and in fact the largest closeness centrality happens at Y.

Next is betweenness. The problem is to find out which node P has the largest betweenness centrality $B(P)$. Refer back to the definition: for any two nodes Q, R, both different and both different from P, we first consider how many shortest (directed!) paths join Q to R, and call this count $S(Q, R)$. There are a (probably smaller) number $S_P(Q, R)$ that go through P. Letting Q, R vary, we sum the ratios $S_P(Q, R)/S(Q, R)$ to get $B(P)$.

Now, since we are considering directed paths, we care about the order of Q and R – for instance, $S(A, Y) = 1$, since there is a single directed edge from A to Y, but $S(Y, A) = 4$, since shortest paths from Y to A have length 3 and go via X_2 after going through either X_1 or X_3, or do the same but with Zs instead of Xs.

So, to compute, say, $B(A)$, we might make a table whose rows are labelled X_1, X_2, X_3, Y, Z_1, Z_2, Z_3 – every node but A – and whose columns are labelled the same way. So, each cell in the table corresponds to a pair (Q, R) of nodes other than A, and in that cell, we put the proportion of the shortest paths from Q to R that go through A (and put, say, a $\times$ in the cells (Q, Q), which are irrelevant because of the definition of betweenness). Then we just sum all the entries in the table (ignoring $\times$) to get $B(A)$.

This is less arduous if we make some observations by looking at Figure 5.1! For instance, to get from any X_i to any Z_j, one has to go through the directed edge from A to Y, and hence through both A and Y. Same, because of symmetry in the network, for going from Z_j to X_i. So all the entries (X_i, Z_j) and (Z_j, X_i) in the table will have a 1.

From Y, you can get anywhere without using A, and it's never faster to use A, so entries that look like $(Y, \bullet)$ all have a 0.

To get from X_1, X_2 or X_3 to either X_1 or X_3, the only way is again through A and Y, but you can get from X_1 or X_3 to X_2 without using either.

We encourage the reader to verify, using observations like these, that A has the highest betweenness centrality, i.e. $B(A) > B(P)$ for all nodes $P \neq A$.

So, we have already illustrated the second caveat:

- from the point of view of in-degree, there's not a clear winner: X_2, A, Z_2 are all "most central", with 2;
- from the point of view of out-degree, Y wins with 4;
- the highest closeness belongs also to Y, with $\frac{1}{11}$, but
- the highest betweenness belongs to A, with 32.

Now let's consider PageRank. To organise things, let's first represent the network as a matrix, called the *(directed) adjacency matrix*. This is, again, just a table with rows and columns labelled by the nodes. In the column corresponding to node P, and the row corresponding to node Q, we put a 1 if there is an edge **from** P **to** Q, and a 0 otherwise:

$$M = \left(\begin{array}{c|cccccccc} & A & X_1 & X_2 & X_3 & Y & Z_1 & Z_2 & Z_3 \\ \hline A & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ X_1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ X_2 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ X_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ Y & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ Z_1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ Z_2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ Z_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right)$$

You can read the in-degrees and out-degrees from the adjacency matrix like this: for each row, the corresponding node has in-degree equal to the sum of the entries in the row. For out-degree, use columns. For instance, the Y row sums to 1, reflecting that $indeg(Y) = 1$, and the Y column sums to 4, reflecting $outdeg(Y) = 4$, both of which are consistent with Figure 5.1.

Here are the out-degrees:

$$outdeg = \left(\begin{array}{cccccccc} A & X_1 & X_2 & X_3 & Y & Z_1 & Z_2 & Z_3 \\ \hline 1 & 1 & 1 & 1 & 4 & 1 & 1 & 1 \end{array}\right)$$

For each node P, we'll write $p(P)$ to mean the PageRank importance score for P. Recall that we solve a certain system of equations to find the various $p(P)$. To be systematic about it, let's define a vector

$$\vec{p} = \begin{pmatrix} p(A) \\ p(X_1) \\ p(X_2) \\ p(X_3) \\ p(Y) \\ p(Z_1) \\ p(Z_2) \\ p(Z_3) \end{pmatrix}$$

From the definition, recall that, say, $p(A)$ is equal to the weighted sum of the PageRank scores of the nodes contributing to the in-degree of A, where the weights are the reciprocals of the out-degrees. The edges coming into A originate at X_2 and Z_2, as can be seen by inspecting the A–row of M for nonzero entries, or looking at Figure 5.1. And X_2, Z_2 both have out-degree 1. So one of our relations is

$$p(A) = p(X_2) + p(Z_2).$$

More generally, we can re-scale the entries in M by replacing each 1 with $1/outdeg(P)$, where P is the node for the column we're looking at. This gives us the *transition matrix*

$$T = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \end{pmatrix}$$

With this, we can package all of the relations between the various PageRanks, coming from the definition, into the single matrix equation

$$T \cdot \vec{p} = \vec{p}.$$

Before we solve it, let's observe that if there is some vector $\vec{p}$ of positive numbers satisfying the above equation, then the same is true for $a\vec{p}$, for any $a > 0$. So the solution won't be unique. But this type of "non-uniqueness by rescaling" doesn't bother us, since the solutions $\vec{p}$ and, say, $2\vec{p}$ give us the same ranking of the nodes by PageRank score: $p(A)$ is greater than, say, $p(Z_1)$ if and only if $2p(A)$ is greater than $2p(Z_1)$.

But that's it. In fact, a positive real solution exists, and it's unique, once you account for non-unqiueness by rescaling.

First, you might have observed that all of the columns now sum to 1, since they previously summed to the out-degree of the corresponding node, and that's what we've rescaled by when changing M into T. This means that T is what's called a *(left) stochastic matrix*. Also, the network is, as a directed graph, connected: any two points, as seen when computing closeness, can be joined by a directed path. This means that the matrix M, and hence T, is what's called an *irreducible matrix*. And then the *Perron-Frobenius theorem*[23] alluded to earlier tells us that a positive vector $\vec{p}$ exists with $T\vec{p} = \vec{p}$.

Here's a more intuitive explanation. Imagine you're looking for me in the network; at any moment, I'm at some node. Since you don't know where I am, you assign, based on whatever information you have, a probability $p_0(A)$ that I'm at A, $p_0(X_1)$ that I'm at X_1, and

[23]The reader unfamiliar with the Perron-Frobenius theorem is **strongly** advised to ignore this paragraph, because we're about to give a more reasonable, down-to-earth, if slightly longer, explanation. In math, there's often the option of giving a long explanation whose individual parts are pretty simple, albeit assembled in a specific way, or a short explanation where many simple parts are hidden in a sealed package with a jargony name, so that the resulting explanation is (1) seemingly impenetrable or complicated and (2) happens as fast as possible. When described that way, the second style of explanation looks obviously bad and the mathematical cultural tendency for preferring brevity over actual simplicity and clarity starts to look bizarre.

so forth, so that all the probabilities (being probabilities) are non-negative and they sum to 1. Let $\vec{p}_0$ be the column vector with entries $p_0(A), p_0(X_1)$, etc.

Then I start to move around the network in the following random way: if I'm at some node P, there are $outdeg(P)$ edges leaving P, and I pick one at random, giving each one a $1/outdeg(P)$ probability of being picked, and move along whichever edge I choose, to its terminal node.

In most cases, I don't have any options: at X_1, there's only one outgoing edge, to X_2, so if I'm at X_1, I move to X_2 with probability 1. It's only when I'm at Y where I have a choice, and here I roll a 4-sided die and go to each of X_1, X_3, Z_1, Z_3 with probability $1/4$.

From your point of view, after one turn, what's the probability that I'm at a given node? Well, to be at A, I must have previously been at either X_2 or Z_2. In either case, I moved to A with probability 1, because I had no choice. Letting $p_1(A)$ be the probability that I'm at A after one roll of the dice, your attitude is therefore something like $p_1(A) = p_0(X_2) + p_0(Z_2)$. What about the probability I'm at X_1? Well, the only way to have arrived at X_1 is if I was at Y before, which had probability $p_0(Y)$, but if I was at Y, then I had four choices for where to go, I gave them an equal chance, and one of them was X_1, so $p_1(X_1) = \frac{1}{4}p_0(Y)$.

In general, $\vec{p}_n = T \cdot \vec{p}_{n-1}$, where $\vec{p}_n$ is the probability vector recording how likely you think I am to be at each node after n iterations of the procedure.

This looks a bit like the equation $\vec{p} = T \cdot \vec{p}$ we want to solve!

By substituting $\vec{p}_{n-1}$ with $T\vec{p}_{n-2}$, etc., we get $\vec{p}_n = T^n\vec{p}_0$. Let's also suppose you had no prior information about my location, so you just gave every node the same initial probability, so $p_0(A) = p_0(X_1) = \cdots = \frac{1}{8}$, reflecting that there are eight nodes.

The question is, does $\vec{p}_n$ hit some sort of equilibrium, i.e. when n is very large, does $T^n\vec{p}_0$ not change very much when we make n still larger? Intuitively, it makes sense that this is pretty similar to the question about solving for $\vec{p}$. Indeed, a solution satisfies $T\vec{p} = \vec{p}$, while p_n tending to some equilibrium means that $T\vec{p}_n$ is *very close* to $\vec{p}_n$, so $\vec{p}_n$ is (in a sense quantifiable using limits) close to being a solution.

And this is what turns out to happen. Roughly speaking: the fact that the columns in T sum to 1 stops the "size" of $T^n\vec{p}_0$ from becoming larger and larger as we increase n. And the fact that T was irreducible stops $T^n\vec{p}_0$ from oscillating between a few separate values as we increase n. That's what we need to ensure that $\vec{p}_n$ converges to some limit, and that's our $\vec{p}$.

It's *easier* to do the linear algebra and solve for $\vec{p}$ than to compute it as a limit of the $\vec{p}_n$, but the latter viewpoint gives an explanation of *why* there is a solution in a way that's relevant to the practical question of how to intepret PageRank. Just solving for $\vec{p}$ tells us that $p(A) = p(Y) = \frac{1}{4}$, and these are the maxima. So PageRank favours A and Y over the other nodes, but doesn't distinguish between them.

Here's the takeaway. If this were a network depicting an actual situation, we'd have to consider the broader context in order to decide which notion of centrality we actually care about, because they give different answers for which node is most central: closeness and out-degree favour Y, betweenness favours A, in-degree favours X_2, A, Z_2 equally, and PageRank favours A and Y equally. We need to have some independent theoretical narrative that ties one of the centrality measures back to the situation we're modelling; a centrality measure without this feature isn't telling us much.

In particular, if we think of the network as more or less fixed in time, and we want to decide which node to "disrupt" in order to make it difficult to move around the network, we'd have to think a bit more to decide whether to disrupt A or Y – is closeness or betweenness more relevant? On the other hand, if we think of the network from the point of view of a particular actor (e.g., a particular dollar in a network of financial relationships) moving around the network, PageRank tells us that we should expect that actor to visit A and Y

most often, and so perhaps we are most interested in those nodes.

6 What does this mean for organizers on the ground?

If a higher centrality measure doesn't automatically imply a higher impact, how can these analyses be put to genuine use in the context of a campaign? We explore this directly in the setting of the Mapping Project and divestment campaigns at the end of the next section, but for now, we introduce some hypothetical situations in order to demonstrate how one centrality measure may be more useful than another depending on what the organizing goals are.

6.1 Scenario A: The home stretch, or the snatch and grab.

Suppose that your community group is low on resources, and also that group members have determined that the end of the struggle is near. Just to make the scenario feel more concrete, imagine that the organizing goal is to stop the development of an oil pipeline through Indigenous land. Perhaps public opinion has truly turned towards the group and against the pipeline; maybe a series of high-profile direct actions have led to a deluge of negative press for the pipeline and community members can sense that they're close to victory.

The group has engaged in some quantitative power mapping: it has compiled data on many of the players involved (eg the oil company, the banks who are financing the project, local and national politicians, etc.) and understands how these entities interconnect. With this data, the group creates a network where the nodes represent potential targets, and the edges correspond to financial support between targets.

Given the facts on the ground, the group realizes that it may not need to plan to disrupt the network for years or decades. One more high-profile hit could be the last straw they're looking for, and given the limitations on resources, it wouldn't be feasible to launch a very long term plan anyway.

The group computes that JP Morgan Chase has the highest out degree, but Wells Fargo has the highest PageRank score[24]. Under these conditions, it might be more strategic to choose JP Morgan Chase as a primary target. The reason is because degree is more sensitive to the short term behaviors of an entity in the network; it expresses how much money one can expect to flow through a given entity in the near future, because this money will be coming from (or moving to) the entities that are adjacent to it in the network.

Under these conditions, it may also be tactical to target nodes with a high betweenness centrality in that betweenness keeps track of the *shortest* paths resources take through the network. Therefore, if a node with high betweenness begins to malfunction, it can have a significant short-term impact on the overall flow of resources through the network.

6.2 Scenario B: The long(er) haul.

Imagine now that we are in the same fight as in Scenario A, but perhaps our group has more resources and/or surmises that the struggle will persist for some time. Because PageRank is not really something that can be gleaned by looking at raw data (unlike a high degree), Wells Fargo has been successful in downplaying its role in sustaining the financial interdependencies upon which the pipeline relies. Under these circumstances, it may be more strategic to the group to choose Wells Fargo as a primary target.

[24]We chose these two banks because both helped to finance the Dakota Access Pipeline.

First, there is the narrative goal of shining a spotlight on a bank that is attempting to minimize public knowledge of its involvement. But beyond this, PageRank ultimately reflects the long-term behavior of a network. Very loosely speaking, if Wells Fargo has the highest PageRank, it means that if we let money flow around the network for a very long time, and only then begin to monitor the activity of any given dollar, we will notice that that dollar has a higher chance of passing through Wells Fargo in a given amount of time than it does any other entity in the network[25]. This interpretation comes from the fact that the PageRank score of a given entity corresponds to a component of a 1-eigenvector for the matrix of coefficients encoding the linear system of importance score equations. Since 1 is the Perron-Frobenius eigenvalue, the long-term behavior of the network is captured by the components of the corresponding eigenvector. So, if the fight is long and if the community group has the resources to maintain the struggle over that time period, PageRank may be a useful centrality tool.

7 A proof of concept.

We now turn our attention to the network compiled by the Mapping Project. Specifically, we will focus on the network whose edges represent financial relationships: there is a directed edge from entity A to entity B when A financially supports B with funds.

Before we begin, we want to emphasize a point that will hopefully by now feel very in line with other disclaimers and cautions that have been mentioned already: **any conclusions drawn from network analysis are dependent upon the full set of choices made when compiling the information to build the network in the first place.** The Mapping Project would have had to make many decisions about which entities to include in the data and which to omit, and in any case there may be important financial relationships that exist between entities that are not publicly known. For these reasons, it's important not to oversell the usefulness of these analyses; it would be most likely impossible to use them to prove beyond a shadow of a doubt that some given entity is objectively "the most important" player in a given collection of real-world social or financial relationships. But at the end of this section, we will describe how a community group might use the sort of quantitative power mapping we demonstrate below to support its strategic choices.

7.1 A (very) basic summary.

Mathematica has many built-in tools for performing the sorts of network analyses we outlined in the previous section. To emphasize how simple it is to start experimenting with these sorts of tools, we intentionally avoided doing anything too sophisticated. Here is a summary of our process:

1. We begin by recording the network in a way that Mathematica will identify. This uses the "DirectedGraph" command in Mathematica: we simply enter all of the information of the network as a list. Each item on the list is of the form "$A \mapsto B$", which means that there is an edge pointing away from A and towards B. See Figure 7.1 below– we

[25] We urge the reader not to take this interpretation too literally. The situation is complicated by the fact that different financial relationships correspond to different amounts of money. So to capture the situation more accurately, one might want to use a version of PageRank that is suited to *weighted* networks– see the subsection entitled "A brief aside on weights" below for more on this. Moreover, even in the setting where one takes weights into account, in order to argue that PageRank accurately reflects the long-term dynamics of the network, one would have to know that the dollar amounts flowing between institutions are relatively stable over long periods of time.

chose names for each entity in the Mapping Project data– you can pick them however you like.

2. The DirectedGraph command will automatically output pictures of the inputted network. Figure 7.1 below shows what one gets with the Mapping Project Data. Already, we have some interesting information that might not have been apparent from staring at the data alone. For example, there are nine distinct *components* of the network; a pair of entities are in the same component exactly when it is possible to start at one and arrive at the other by traveling along edges. So for example, if the main goal of an organizing group is to impact entities in a given component, it *may* not make much sense to dedicate lots of time and focus on entities in an entirely different component. We emphasize the "may" in the previous sentence to remind the reader of the disclaimer at the beginning of the section: targeting entity A in order to ultimately impact entity B could be a reasonable move even if A and B lie in different components, if for example A and B are connected in the real world in ways that the data used to create the network do not reflect.
3. We now use built-in tools for computing centrality measures: "BetweennessCentrality", "ClosenessCentrality", and "PageRankCentrality". For those unfamiliar with Mathematica, it's as simple as plugging the list you already created to represent the network into those commands. Mathematica will then output a list of centrality scores– one for each entity in the network. We therefore compute three different centrality scores for each entity: a betweenness score, a closeness score, and a page rank score. Note that we don't bother computing the degree score, since it's very easy to access from the raw data: an entity has a high degree exactly when it appears many times in the data as an organization that supports many other organizations. For the record, the entity with the highest degree in the Mapping Project data is Combined Jewish Philanthropies (CJP), a non-profit that pools funds together from multiple donors to support various causes. They partner closely with the Anti-Defamation League (see our footnote above for more info on the ADL) and characterize its efforts to combat anti-zionism as one of its most important projects[26].

[26] To be clear, CJP refers to its work as "fighting anti-semitism and anti-Zionism", making clear that they view the latter as a dangerous example of the former.

```
In[101]:= DirectedGraph[{AstraZeneca → MassGenHospital, Berklee → Baupost,
    GatesMedResearch → NewEnglandInfectious, GatesMedicalResearch → EditasMed,
    Biogen → MassGenHospital, BostonPoliceDeptFoundation → BostonPoliceDept,
    Brandeis → Baupost, CentralRegionHomelandSecurity → CentralMassLawEnforcement,
    CombinedJewishPhil → ACLU, CombinedJewishPhil → DavidProject,
    CombinedJewishPhil → Berklee, CombinedJewishPhil → HarvardBusiness,
    CombinedJewishPhil → JStreetBoston, CombinedJewishPhil → HarvardMed,
    CombinedJewishPhil → BostonCollege, CombinedJewishPhil → JewishNatlFundNE,
    CombinedJewishPhil → BroadInstitute, CombinedJewishPhil → BostonMuseumScience,
    CombinedJewishPhil → Emerson, CombinedJewishPhil → Lesley,
    CombinedJewishPhil → CAMERA, CombinedJewishPhil → Suffolk,
    CombinedJewishPhil → TuftsMedSchool, CombinedJewishPhil → TuftsMedCenter,
    CombinedJewishPhil → Wellesley, CombinedJewishPhil → BostonPoliceDeptFoundation,
    CombinedJewishPhil → Clark, CombinedJewishPhil → AmericanJewishCommitteeNE,
    CombinedJewishPhil → NewIsraelFundBoston, CombinedJewishPhil → Gann,
    CombinedJewishPhil → Tufts, CombinedJewishPhil → UMassAmherst,
    CombinedJewishPhil → BostonU, CombinedJewishPhil → NortheasternU,
    CombinedJewishPhil → SynagogueCouncilMass, CombinedJewishPhil → HarvardU,
    CombinedJewishPhil → JewishBoston, CombinedJewishPhil → Lappin,
    CombinedJewishPhil → MIT, CombinedJewishPhil → FriendsIDFne,
    CombinedJewishPhil → JewishJournal, CombinedJewishPhil → LIBI,
    CombinedJewishPhil → SalemState, CombinedJewishPhil → ADL,
    CombinedJewishPhil → FacingHistory, CombinedJewishPhil → NEJewishLabor,
    CombinedJewishPhil → JewishTeenBoston, CombinedJewishPhil → Brandeis,
```

Figure 2: What it looks like when inputting a directed graph into Mathematica. Since the entire network has many entities, we cut the image off – the figure shows only a portion of the edges in the network.

7.2 A brief aside on weights.

We have not discussed the notion of a *weighted* network in this piece; this is a network with some extra information consisting of an assignment of a (usually) positive number to each edge. In the context of the Mapping Project, we might consider assigning an edge with the amount of financial support (measured in US dollars) flowing from entity A to entity B.

One could imagine a financial relationship network in which a given entity, A, has a very high centrality score (using for example betweenness, closeness, or PageRank). But this may not mean much in terms of the actual impact A has on the entire network if every edge involving A corresponds to very small amounts of money. Conversely, there could be another entity, B, which has a very low centrality score– to imagine an extreme situation, perhaps B connects only to one other entity and no others. But if B is pumping this one other entity with gigantic amounts of money, and then *that* entity passes it along to many other nodes in the network, the real-world impact of B will not be reflected by its poor centrality.

For these reasons, it can often be worthwhile to keep track of weights when that information is available. There are then versions of centrality measures that are designed to work in this more general, weighted context. To keep matters simple[27] and because our only goal is to demonstrate the potential of these analyses, we don't investigate this possibility further here. But it's worth mentioning that these techniques exist, and so we do.

Intuitively, it makes more sense to keep track of weights when they vary greatly across the network. In other words, if it were true in a financial relationship network that every entity outputs and inputs approximately the same amounts of money, we can be more

[27] And also, because exact financial amounts aren't always readily available in the Mapping Project Data.

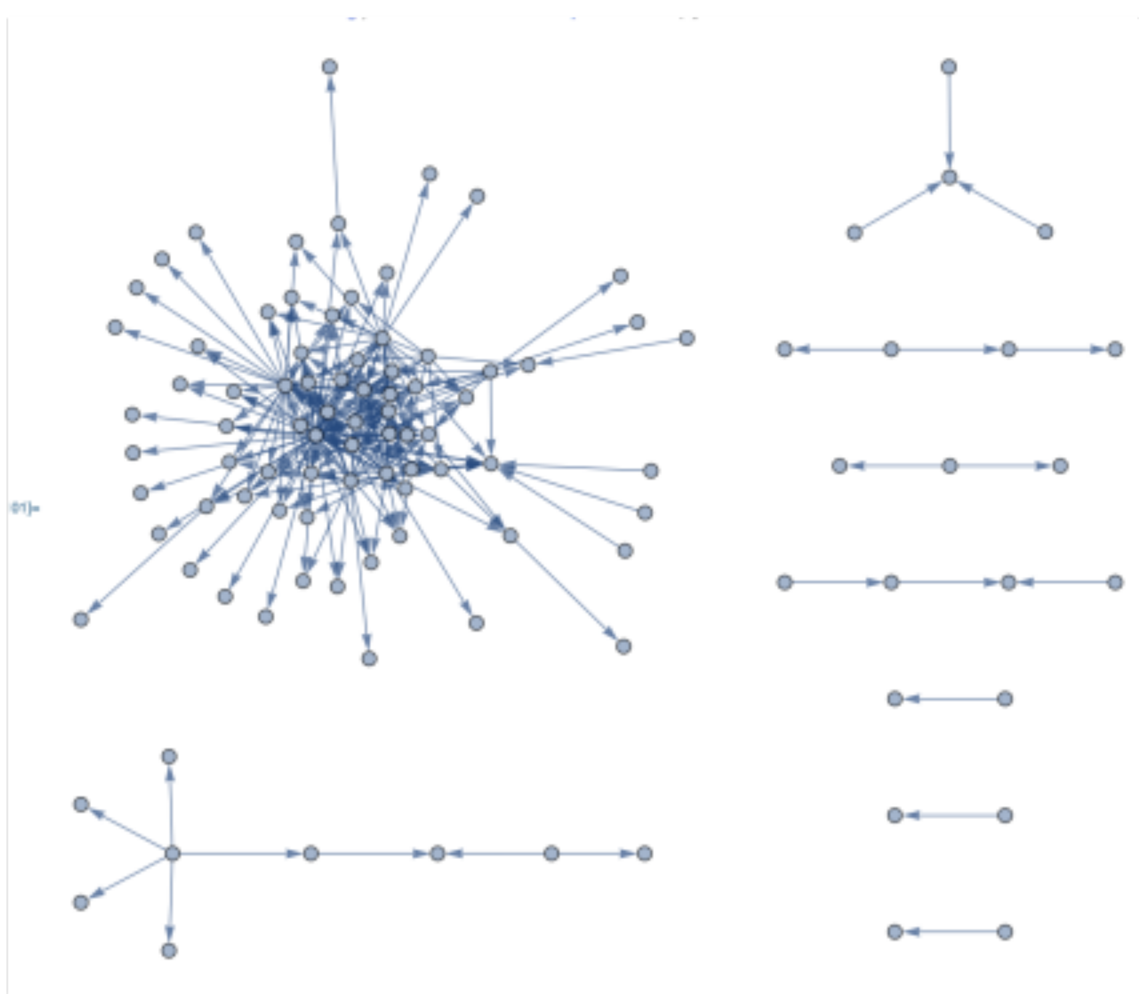

Figure 3: A schematic of the network that Mathematica automatically outputs when a user inputs a directed network as in the Figure above.

confident that unweighted centrality measures are capturing roughly the same information that we would learn by keeping track of weights.

On the other hand, even when this isn't the case, a genuinely thorough analysis might include computations of both weighted and unweighted centrality measures. That is to say: there is often something interesting to learn from the unweighted centrality measures, even when the amounts of money flowing out of or into different nodes varies greatly across the network. Unweighted centrality tells us about the "shape" of the network. For example, imagine a node A within a financial network which when deleted, cuts the network into two disconnected pieces. Even if the money flowing directly into or out of A is relatively small compared to amounts occurring elsewhere, A plays a very important role in the network in that all money moving from one side of the network to the other must pass through it, and unweighted centrality measures will detect these sorts of properties.

7.3 Findings.

In this section, we briefly summarize the outputs of the very basic analysis described above.

- For **betweenness centrality**, far and away the entity with the largest score is CJP (mentioned above in the subsection summarizing the Mathematica procedure) with a score of 275. The next highest betweenness score of 32 belongs to Fidelity Charitable, and then comes the Boston-based hedge fund Baupost Group with a score of 13.
- For **closeness centrality**, there are several entities tied for the largest score (recorded in no particular order): AstraZeneca, Baupost Group, the Bill and Melinda Gates Medical Research Institute, Biogen, the Boston Police Department Foundation (not to be confused with the Boston Police Department itself), Millenium/Takeda, Central Regional Homeland Security Advisory Council, Southeast Regional Homeland Security Advisory Council, Northeast Homeland Security Advisory Council, the Broad Institute, L3Harris, Raytheon, Fidelity Charitable, IBM, Google, Merck, and the State of Massachusetts.

  The fact that so many entities tied for the largest closeness centrality is related to a deeper geometric truth of the network: there just aren't that many different numbers for which some pair of nodes has that number as their distance. If the distance from some node to some other node even exists in the first place (remember that there are pairs of nodes that live in different components, meaning there is just no way to go from one to the other and in this instance, there is no notion of distance between them), their distance is either 1, 2, or 3 (and is most probably either 1 or 2). Since there aren't that many options for distance, there in turn aren't going to be many options for average distance, and so there aren't many closeness centrality scores that arise. This suggests that closeness centrality really strips away a lot of the complexity of the relationships in the network, and therefore greater care should be taken before drawing serious conclusions from it.

- For **Page Rank**[28], the entity with the highest score is (easily) Harvard University; second place is Baupost Group; and third is the Massachusetts General Hospital system[29]. It's a good time to refer back to the disclaimer at the beginning of this section: results will obviously depend sensitively on whatever choices the data collectors make in designing the network. And so perhaps we shouldn't be surprised to see Harvard achieving such a high page rank score given what the Mapping Project says about its own work, namely that they have "emphasized the university as a central nexus that ties together many of the harms traced on the map." We will delve further into this point in the next subsection.

7.4 Using quantitative power mapping in a campaign.

How can the sorts of findings that we outline in the previous subsection be put to use by a community group, hoping to hone its strategy or justify its choice of target? This question is especially warranted considering once more the disclaimer from the beginning of this section: the conclusions we draw are a function of the choices data collectors make when compiling the information that gives rise to the network, so perhaps what we are really learning with our analysis is much more about those choices than about real-world dynamics.

A skeptic could easily argue– as we suggest at the end of the last subsection– that of course Harvard University achieves a high Page Rank score: the Mapping Project itself makes clear that it chose to emphasize the role that major universities in Boston play in sustaining the forces of Zionism and settler-colonialism. So, the skeptic continues, either:

1. Harvard is not– in any "real" sense– at the center of it all; what we see with Page Rank analysis is just a reflection of the collection of arbitrary choices made by the data compilers to manufacture a network in which Harvard plays a central role; or
2. Harvard *is* in some real sense at the center of it all, but then evidently, we didn't need the network analysis to deduce this because after all, the Mapping Project already

[28] We note that the most basic version of PageRank assumes that the network is connected. This corresponds to the matrix of coefficients being irreducible. However, Figure 3 makes clear that the Mapping Project financial network is not connected. In this case, the matrix of coefficients decomposes into a block matrix and one performs PageRank on each block separately.

[29] The reader might be surprised to see a hospital appearing in these data. We refer to the Mapping Project which has already done a great job of justifying their inclusion: "According to government spending records, MGH has received over 100 million in funding from US Department of Defense (source: USASpending.gov), and maintains numerous links to the Pentagon. MGH has hosted a series of "CEO Firesides" that feature both corporate executives (from companies such as Novartis) and representatives of the US national security state. One MGH Fireside Chat featured General Keith Alexander, former head of the NSA, along with Biogen CEO Michel Vounatsos and Head of Research and Development at Bayer Pharma Joerg Moeller. In MGH's advertisement for this panel, Alexander was described as "Former head of NSA, retired four star Army General and global security strategist." Like other high-ups in the US national security state, Alexander used his background to become an entrepreneur and "security consultant" to corporate America. Alexander was appointed to Amazon's Board of Directors after the company had trouble securing Pentagon contracts. MGH's vaccine development and pandemic preparedness research centers around its desire to receive funding from the Pentagon."

knew that universities are responsible for propping up these sorts of networks which is why they made the choices they did in the first place.

In fact, we agree with the skeptic! And in particular, we're of the mind that it's the second of these two possibilities that is actually occurring: Harvard is, in *some* sense, really playing a central role in maintaining these systems of oppression, and network analysis isn't necessary for concluding this. Indeed, we are not (necessarily) proposing quantitative power mapping as a diagnostic tool, that is, as a means of determining what is actually going on. On the contrary, we believe that people with organizing experience often have much more reliable and intuitive ways of knowing the facts on the ground and who the major players are, just by virtue of living alertly in the real world. We have both experienced firsthand being parts of community-led movements that managed to reach consensus on the appropriate targets for the struggle at hand, without using any high-tech mathematics or computational techniques.

This presents the question: if quantitative power mapping isn't meant for determining who to target, then what's it good for? **It is a way to leverage the social power that mathematics commands, to support and defend the choices that a community group already knows are the strategically correct ones to make.** Let's explore this point further in the context of Harvard and the Mapping Project.

Harvard has a multi-decade history of attempting to argue that the investments of its endowment fund should be seen as apolitical; the university simply makes the investment decisions that are conventional in the finance world. And as the argument goes, this makes perfect sense, since after all Harvard and the work that goes on there to better understand our world is so essential, and the higher the endowment gets the more easily the university can subsidize tuition for poorer students. A number of more specific arguments have been used by Harvard administrators to avoid financial accountability over the years, for example:

- In the mid '80's, President Derek Bok (`https://www.thecrimson.com/article/1984/10/2/the-problem-of-divestment-pfollowing-is/`) argued against divesting from Apartheid South Africa, claiming that *first and foremost*, we can all agree that what the South African apartheid government is doing is wrong and bad, but divestment will cost the university money; that anyway Harvard is doing its part by voting in non-binding resolutions urging private companies that operate in South Africa to adopt minimal ethical standards, and also that Harvard has a program for recruiting Black South African students; that it's just not right for a university to impose its political views on external bodies and that doing so would corrupt the university's position as a cloistered apolitical realm of pure thought exploration; (somewhat paradoxically given the last argument) divestment implies ceding all power Harvard has to influence companies that operate in South Africa; that if Harvard chooses to divest, it opens itself up to other divestment campaigns from people who might disagree with its own politics; and finally, even if none of those arguments are convincing, that divestment won't successfully impact the situation. Note that Harvard did in fact choose to divest two years after this letter appeared.
- In an official press statement in 2013 (following the recent founding of the group Divest Harvard which aimed to organize community members towards putting sufficient pressure on administration to divest from the fossil fuel industry), President Drew Faust argued against divesting from fossil fuels (`https://www.harvard.edu/president/news-faust/2013/fossil-fuel-divestment-statement/`). *First and foremost*, he makes clear that we can all agree that climate change is scary and that it is wrong and bad to continue to consume fossil fuels at current rates, but that Harvard exists primarily to serve an "academic mission" and not a political one; (somewhat

paradoxically given the last argument) that to divest would amount to Harvard ceding all power to influence big companies to act responsibly when it comes to fossil fuels; that Harvard is doing its part by offering lots of courses on climate change and that Harvard even gives out special awards for acting sustainably; that divesting could be offensive to donors who may choose not to donate further; that anyway universities account for a very small portion of the fossil fuel investment market, and if Harvard divested, others would just pick up the dropped investments so in the end it's not even an effective tactic. Note that as of 2021, Harvard did choose to divest from fossil fuels.

- When it comes to Palestine and the question of divestment in the spirit of the BDS movement, it seems that campus culture amongst high powered administrators hasn't yet evolved to the point of the analog of the "first and foremost" assertions above: there is nothing near consensus amongst university leaders on the nature of Israel's actions and whether or not they constitute apartheid, ethnic cleansing, or settler-colonialism. So, the administrative response to student-led initiatives in favor of BDS attacks these grassroot efforts on far less subtle grounds. There is thus as of yet no need to make arguments along the lines of calling efficacy of BDS into question, or arguing for the importance of remaining politically unbiased. We refer to the Harvard Crimson editorial staff recently endorsing BDS in 2022[30] and, for instance, ex-president Larry Summers' response piece in the New York Sun [31].

As the tide continues to turn– and it will– towards the realities on the ground in Palestine and more people realize the situation for what it is, one will expect university administrations to resist change in much the same ways that they've done in the past– by arguing that the university is really not involved in politics and that in any case its actions account for a tiny fraction of the problem at hand. For these reasons, we argue that a mathematical analysis which can point to –and in fact in some sense quantify– Harvard's centrality to maintaining the problem, could be a very useful tool for future divestment campaigners. The skeptic can attempt to argue that choices were made in data compilation to artificially place Harvard at the center of things, but they can not successfully argue that there is no way to see or to conceive of Harvard as being centrally important to the maintenance of these forces, as it has claimed time and time again before. The findings above do not literally assert the following statement, but in *spirit* the high page rank score of Harvard can be interpreted as follows (see the footnote for a clarification): as a dollar flows through the network compiled by the Mapping Project– a network that undeniably facilitates Zionism and settler-colonialism – it is highly likely to pass through Harvard's coin purse in a set amount of time[32]. And while there may have been other ways to compile the network, or they may exist altogether different networks operating in the Boston area, administrators would have to argue for the overall non-importance of this *particular* network– comprised of over 250 entities shuffling

[30] See `https://www.thecrimson.com/article/1984/10/2/the-problem-of-divestment-pfollowing-is/`.

[31] Summers: `https://www.nysun.com/article/antisemitism-comes-to-harvard-in-both-intent-and-effect`.

[32] This statement should be taken with a big grain of salt, primarily because we did not take edge weights into consideration. The high page rank score for Harvard is based on the geometry of the network alone, and not on the actual amounts of money flowing from any one entity to any other. So, this statement could be taken more literally if it were true that every financial relationship in the network corresponded to the same dollar amount, which of course is not the case. However, we emphasize that the actual dollar amounts flowing in and out of Harvard that are recorded in the data are gargantuan: at least half a billion dollars coming in and over 2 billion dollars out (much of that in the form of investment commitments). We refer the reader to the more general measure known as TextRank which can take edge weights into account: `https://digital.library.unt.edu/ark:/67531/metadc30962/m2/1/high_res_d/Mihalcea-2004-TextRank-Bringing_Order_into_Texts.pdf`. We also emphasize that financial relationships can wax and wane, or pop in and out of existence in time; dollar amounts can change drastically as new donors enter the picture. Nevertheless, the data compiled by the Mapping Project provides an informative snapshot of a network that evolves through time.

billions of dollars back and forth between them– to make the same sorts of claims it's made in the past regarding Harvard's actions being inconsequential.

8 Conclusion and summary

We have argued that given a network of entities connected to one another by for instance financial dependencies, computing centrality measures can be a useful way for an organization to justify its choices for how to target specific players. We outlined the basics for several centrality measures (degree, betweenness, closeness, and PageRank) and computed all of them for the financial network compiled by the Mapping Project. We then described how a grassroots campaign might use these sorts of analyses in fighting for major universities to divest from Israel until or unless the goals of the BDS movement have been met.

We emphasize once more that we barely scratch the surface. We don't perform any analysis on weighted networks– let alone time-dependent networks where one allows for weights to evolve as functions of time– and it would be interesting to gather the necessary information to run these more sophisticated calculations.

We hope that at the very least, this piece inspires thought amongst mathematically trained people, and encourages them to wonder how their training can be useful in movement spaces. There is more to being a politically conscious mathematician than feeling guilt for how the field has been put to use by militaries, reactionary forces, and authoritarian governments. We have both the power and the responsibility to imagine how mathematics fits into a free and just world.