Characterising the structure of the largest online commercial sex network in the UK: observational study with implications for STI prevention

Authors

Luca Giommoni[1*], Giulia Berlusconi[2], and G.J. Melendez-Torres[3]

1. School of Social Sciences, Cardiff University, UK

2. Department of Sociology, University of Surrey, UK

3. Peninsula Technology Assessment Group, College of Medicine and Health, University of Exeter, UK

*Address: Cardiff University, Glamorgan Building, King Edward VII Avenue, Cardiff CF10 3WT. Email: giommonil@cardiff.ac.uk. Phone: +44 2920874436

Word count: 7,498

Abstract

This study analyses large-scale online data to examine the characteristics of a national commercial sex network of off-street female sex workers and their male clients to draw implications for public health policy and practice. We collected sexual contact information from the largest online community dedicated to reviewing sex workers' services in the UK. We built the sexual network using reviews reported between January 2014 and December 2017. We then quantified network parameters using social network analysis measures. The network is composed of 6477 vertices with 59% of them concentred in a 'giant component' clustered around London and Milton Keynes. We

found minimal disassortative mixing by degree between sex workers and their clients, and that a few clients and sex workers are highly connected whilst the majority only have one or few sexual contacts. Finally, our simulation models suggested that prevention strategies targeting both sex workers and clients with high centrality scores were the most effective in reducing network connectedness and average closeness centrality scores, thus limiting transmission of STIs.

Key words: commercial sex network; large-scale online data; social network analysis

Introduction

Sex workers and their clients remain at high risk of contracting sexually transmitted infections (STIs) (Dias 2015; Mc Grath-Lone et al. 2014). Indeed, commercial sex work represents an important channel for the diffusion of STIs (Shannon et al. 2014), and their prevalence is estimated to be high among female sex workers in European countries whilst HIV prevalence is generally low (Platt et al. 2011). This also appears to be the case in the UK, where STIs tend to be more prevalent among both female sex workers and men paying for sex, although an increasing diversity of sex markets has led to different risk levels of STI infection (Dias 2015). Clients remain at greater risk of acquiring STIs and contributing to their transmission (Jones et al. 2015). A study of indoor-working female sex workers in London found that migrant sex workers tend to see more clients and are less likely to use contraception than UK-born ones, although both groups reported more consistent condom use for penetrative sex than oral sex (Platt et al. 2011). There remains, however, little recent epidemiological knowledge as to the specific protective behaviours used by female sex workers in the UK or European contexts.

The spread of infectious diseases across networks can be modelled modelled as simple contagion, i.e. contagion 'for which a single activated source can be sufficient for transmission' (Guilbeault, Becker, and Centola 2018, 4). Effective STI prevention requires understanding the structure and composition of the sexual networks across which infections are transmitted through simple contagion, i.e. the set of individuals and the sexual relationships among them (Chami et al. 2017; Centola 2018). However, obtaining a complete map of commercial sex networks using traditional data collection methods (e.g. contact tracing or census data) presents several challenges (Zhang and Centola 2019; Klovdahl 2005), and studies addressing structural network characteristics usually draw on relatively small, geographically limited populations (Shushtari et al. 2018). Internet-

based sex markets have grown in recent years (Sanders et al. 2018). They offer sex workers and their clients new ways to communicate with each other, and provide new opportunities to develop effective interventions to target large populations and reduce STI diffusion (Hsieh et al. 2014). The analysis presented in this article builds on the literature on sexual networks. It draws on user-generated data from a popular website dedicated to review sex workers' services to create an original empirical dataset of the commercial sex network in the UK, with implications for STI prevention.

Sexual networks and STIs

The study of sexual networks plays a crucial role in understanding both the rate and the extent of STI diffusion (Newman 2002). First, network structural characteristics such as network cohesion, average path and tendency toward clustering, and connectivity can tell us how quickly or how far STIs might spread across communities (Campbell and Salathé 2013). For example, in a highly dense network or a network where individuals have high contact rates, diseases spread quickly as most of the members are closely connected to each other (Doherty et al. 2005). STI spread is also accelerated by relatively short paths between any two individuals, and the tendency toward clustering, i.e. the tendency of an individual's contacts to have contacts among each other and to cluster into densely connected groups (Guilbeault, Becker, and Centola 2018). The presence of many unconnected pairs instead affects the incidence of STIs as it limits the extent of their diffusion (Doherty et al. 2005).

Second, actors' positions within the network increase their risks of contracting STIs and their likelihood of spreading infection (May and Lloyd 2001). Rothenberg et al. (2007) showed, in a community of teenagers in rural Georgia, that participants with syphilis had a higher degree and betweenness centrality than participants without syphilis, where degree centrality measures an actor's number of relationships with other network

participants, and betweenness centrality measures the extent to which an actor lies on the shortest path between any pair of network participants (Wasserman and Faust 1994). The latter, in particular, is useful to identify bridging people, i.e. individuals who are more likely to facilitate STI spread across communities (Youm 2015).

The distribution of centrality scores across network participants is also relevant. Sexual networks are usually characterised by a small number of very active individuals, and a large number of actors with only one or few sexual contacts. A positively skewed degree distribution increases the rate of STI diffusion (Newman 2002) but also makes the immunisation of these highly connected individuals particularly effective (Doherty et al. 2005). Finally, assortative mixing by degree, i.e. the tendency of individuals to interact with others with similar level of activity, increases STI diffusion rate, but limits the extent of spreading (May and Lloyd 2001). Disassortative mixing, on the other hand, occurs when there are contacts between highly connected and less-connected actors, and increases the extent of STI diffusion (Youm 2015).

Despite the value of sexual networks for understanding STI risk, few studies have considered the network structure, position and composition of commercial sex networks (Schrager et al. 2013; Latkin et al. 2011). A systematic review of social network analyses of female sex workers and HIV risk behaviours found only four studies addressing structural network characteristics, each of which drew on relatively small, geographically limited populations (Shushtari et al. 2018). There are two main reasons for these gaps. First, these studies require the use of complete network design, i.e. the collection of data from all the members of a community, which is expensive, time consuming and raises ethical concerns as it asks participants to name their sexual partners (Klovdahl 2005). Second, hidden populations such as sex workers and their clients are typically hard to reach (Valente and Pitts 2017). Thus, prior studies have been small-scale in nature or have

relied on contact tracing to study egocentric networks. Both approaches miss the complexity and heterogeneity of commercial sexual networks at regional and even national levels (Hao et al. 2015; Klovdahl 2005).

Sex work and digital technologies

The internet has had a transformative impact on the sex industry and, consequently, on the way we research it. Digital platforms changed the way sex workers and clients interact before the in-person activity takes place (Sanders et al. 2018; Stewart Cunningham et al. 2018). Before the advent of the internet, sex workers could reach clients by streetwalking specific urban areas, or working in brothels, massage parlours and walk-ups (Scott Cunningham and Kendall 2011; Crotty and Bouché 2018). Today sex workers can advertise their services, and be contacted by clients, via advertising platforms, agency-owned or personal websites, and social media platforms (e.g. Twitter or Facebook) (Grov et al. 2017; Gezinski et al. 2016). Since the beginning of the 2000s, there has also been a large diffusion of customer review websites, where clients can write and share detailed descriptions of their experiences with sex workers (Crotty and Bouché 2018; Gezinski et al. 2016).

These reviews can play an important role in potential clients' decision-making (Sanders et al. 2018). In some cases, a negative review is enough for putting a sex worker out of business. On the contrary, positive reviews can help build trust among clients that the sex worker is genuine and 'professional' (Sanders et al. 2019; Noack-Lundberg et al. 2019). Analysing a Brazilian online community over six years, Rocha et al. (2010) found that a good review is a predictor of the future popularity of the sex worker. This means that reviews from clients can alter sex workers' centrality within commercial sex networks by attracting both many local clients and sex tourists from other cities. The direct consequence of this is that customers' online activities such as forum discussions

and reviews can shape offline interactions between sex workers and their clients. Contemporaneously, the study finds a strong influence of offline factors, such as urbanity and geography, on the network structure. Hsieh et al. (2014) find similar results regarding the relevance of geography from the analysis of online communities in Brazil and the US. The authors find that the travelling of clients and sex workers to different locations can explain around 50% of their centrality in the network.

Therefore, the study of online commercial sex networks is important for at least two reasons. First, given their popularity among both sex workers and clients, these platforms are a valuable source of information for understanding the structure of online communities (Rocha, Liljeros, and Holme 2010; 2011; Hsieh, Kovářík, and Logan 2014). For instance, data from these platforms can be used to identify key players in the virtual community, i.e. popular sex workers and active clients (Zhang and Centola 2019). Second, and perhaps more importantly, as online and offline networks overlap and shape each other, online data can provide an insight into the structure and composition of offline commercial sex networks (Rocha, Liljeros, and Holme 2010; Hsieh, Kovářík, and Logan 2014). Sexual contacts extracted from popular online communities can compensate for the lack of traditional data on complete networks, and can be analysed to suggest prevention strategies based on network properties.

The current study

This study contributes to the research on sexual networks by examining a national commercial sex network of off-street female sex workers (i.e. sex workers in commercial venues) and their male clients. While abundant research exists on both the application of social network analysis for public health interventions and online sex communities, the linkage between these two dimensions is still under-explored. This study uses internet-mediated data as an alternative approach to sequenced sampling to collect large scale

sexual contact data. Specifically, we collected sexual contact information from the largest online community dedicated to reviewing female sex workers' services in the UK. The study builds on previous research on online socio-sexual networks (Hsieh, Kovářík, and Logan 2014; Rocha, Liljeros, and Holme 2011) by answering the following questions. What are the main structural characteristics of this online sex community? How are direct sexual contacts distributed across sex workers and their clients? How can network structural characteristics and individual positions suggest effective STI prevention strategies?

Method

Data and procedure

The online community from which we collected our data is openly accessible to anyone, although visitors need to confirm that they are older than eighteen the first time they access the website. It was created in 1999 for the exchange of information between sex workers and clients. The website specifically focuses on the off-street section of the sex market, which includes both sex workers working independently or for third parties. The off-street sex market represents the largest sector of the sex market in England and Wales, with some figures showing that up to three quarter of sex workers work in various indoor settings (Home Affairs Committee 2016). The platform offers several services such as a message board, escort advertisements, and web camming. However, its main function is reporting male clients' reviews of female sex workers. Each review contains dyadic information about client's username and sex worker's name, date and time, city, venue (e.g. escort agency, massage parlour), duration of the encounter, price paid, and three written accounts describing each of the venue, the sex worker and the intercourse. While

clients need to login into the platform to provide a review, the website is free to search for service providers, reviews and sex worker's profiles. Note that reviews refer exclusively to face-to-face encounters, and not to services such as web camming or instant messaging.

We developed a crawling and scraping software to collect this information from the online community. The software automatically and daily accesses, crawls, fetches and stores this information to a database. We analysed data reported between January 2014 and December 2017.

We identified clients and sex workers by their usernames. Each client and each sex worker formed a vertex in the network. The identification of unique clients was straightforward as this relied on unique account names. However, identification of unique sex workers was more complicated as they were identified by generic, client-reported street names such as *Bethan* or *Cleo*. Our approach to identifying unique sex workers was conservative. We assumed that two or more reviews reporting the same name, for instance *Bethan*, referred to the same sex worker if they also reported the same venue and city. For instance, Review #1 and Review #2 refer to the same sex worker, *Bethan*, if they both report the same venue, i.e. *Venue Y*, and the same city, i.e. London. Two reviews reporting the same name and the same city but different venues, for instance *Venue X* and *Y*, refer instead to two different sex workers.

To check accuracy and validity of our approach to the identification of sex workers, we inspected a random sample of 500 reviews reporting the same names but for different venues and cities to ensure that this approach was valid. Because reviews include descriptions of sex workers' appearance, we used this additional information to understand if, for instance, the *Bethan* in London is the same working in Cardiff. Despite sex workers sometimes working in more than one geographical area regularly or for a

short period of time (Sanders et al. 2018), in none of these checks were we able to identify with confidence if a sex worker was working in two different cities or venues at the same time, suggesting the validity of our conservative approach.

We also performed an additional check that our process did not merge separate sex workers into one. For 200 sex workers with more than 4 reviews we used the information about the sex worker's appearance to check whether they reported significant differences. For instance, descriptions reporting different ethnicity, physique, age, etc. for the same identified sex worker would point out a mistake in our coding process. In all the checks performed, we were confident that, from the information provided, the descriptions referred to the same sex worker, confirming the soundness of our approach.

Finally, we transformed reviews into a map of a national off-street commercial sex network. Specifically, we established a link between *Client A* and *Sex Worker B*, every time *Client A* posted a review about *Sex Worker B*. We abstracted the frequency of each link assuming that once formed the link is persistent; that is, multiple reviews for the same sex worker from the same client formed a single connection. This process resulted in a binary, bipartite network as clients cannot directly connect to other clients, and sex workers cannot directly connect to other sex workers.

The resulting network is likely to be a specific sub-set of the British off-street sex market of female sex workers and their male clients. Mapping the online sex industry in the UK is a challenging endeavour, and data collected from advertising or other online platforms are never complete (Sanders et al. 2018). Reviews are written by a minority of clients (Sanders et al. 2018) who seem to be among the most experienced or active ones. Indeed, clients often refer to previous experiences when writing their reviews, and use the same, unique argot (Holt and Blevins 2007).

Measures

We quantified network parameters using social network analysis measures.[1] First, we identified the main structural characteristics of the commercial sex network which are known to affect STI diffusion (Hsieh, Kovářík, and Logan 2014). Degree assortativity is the tendency of vertices with similar number of links to preferentially associate with each other (Newman 2003). The giant component is the largest connected subset of vertices in the network. For this subset, we can also calculate the average geodesic distance, i.e. the mean shortest path between any two vertices, and the diameter, i.e. the maximum distance between two vertices of the giant component (Wasserman and Faust 1994). We also calculated the clustering coefficient using Opsahl's approach (2011) to detect clustering in a two-mode network.

Second, we calculated three different centrality scores for each vertex (sex workers and clients) in the network: degree centrality, betweenness centrality, and closeness centrality. Degree centrality measures the number of individuals with which each person in the network is connected. Betweenness centrality instead measures the number of times a vertex is along the shortest paths between any two other vertices in the network (Wasserman and Faust 1994). This measure is often used to identify individuals who are central because they are brokers; that is, they control communications or facilitate the exchange of resources within the network. Finally, closeness centrality is the inverse of farness, which measures the distance of an individual from every other individual in the network. We use Opsahl et al.'s (2010) approach, which enables us to calculate a closeness score for each individual in the network despite the commercial sex network having several disconnected components.

---

[1] The analyses are performed using the statnet suite of packages (Handcock et al. 2005) and the tnet package (Opsahl 2009) for R (R Core Team 2019).

In the last part of the analysis we tested the impact of different vertex removal strategies on the level of connectivity of the network, and on individuals' closeness to all other vertices in the network. Connectivity was measured using Krackhardt's connectedness score, which is equal to the fraction of all dyads (namely the combination of two vertices) connected through an undirected path (Krackhardt 1994). Individuals' closeness was measured by calculating the average closeness centrality score for all vertices in the network. By using Opsahl et al.'s (2010) approach, once again we circumvent the issue of having several disconnected components. Vertices were removed 1) randomly, 2) based on degree centrality scores, 3) based on betweenness centrality scores, and 4) based on closeness centrality scores. For random removal, we simulated vertex removal 100 times, and reported the average connectedness score and closeness score for the 100 replications. Following Hsieh et al.'s approach (2014), vertex removal was also based on the role of individuals within the network. Each of the three strategies mentioned above was thus applied to 1) clients, 2) sex workers, and 3) both clients and sex workers. The twelve resulting removal strategies approximate the impacts of different strategies which might focus on individuals with a specific role or position within the network. Although we gradually removed all vertices from the network, following Hsieh et al.'s example (2014), we report connectedness scores and average closeness scores only for the first 5% as we expect STI prevention strategies to reach only a limited number of individuals and to be effective when a small fraction of individuals are targeted (Newman 2002). Finally, it is worth mentioning that for the purpose of this study, targeting vertices does not necessarily mean physically removing people from the network but rather using immunisation strategies to prevent STI transmission.

Results

FIGURE 1 ABOUT HERE

TABLE 1 ABOUT HERE

General network characteristics

The network is composed of 6,426 edges, or connections between sex workers and clients. Because each vertex corresponds to one sex worker or one client, 6,477 people were represented in the network, of which most are sex workers (60%). In addition, the network covers 1,656 venues and 328 geographical locations (see Table 1). Greater London and South East England, with respectively 3,154 and 1,502, are the two geographical areas concentrating the highest number of reviews of commercial sex encounters.

A key feature of networks is the 'giant component', i.e. the largest connected component that includes a considerable proportion of the actors in the network. In this network, the giant component includes 3,807 people (2,342 sex workers and 1,465 clients), concentrating 59% of all people, and is localised around London and the London exurb of Milton Keynes. A total of 2,846 sex workers and clients were active in these two areas between 2014 and 2017 (2,079 in London and 767 in Milton Keynes), and many of them are also part of the giant component. Two chi-square tests confirmed the presence of an association between being in the giant component and being in London (χ^2 (1) = 60.1, p <.01) or Milton Keynes (χ^2 (1) = 419.3, p <.01).

This means that, if we assume that the adoption of safe sex practices is equally distributed in the network, buying or performing sex in these two cities put people at a potential higher risk of contracting STIs, given the higher interconnectivity among vertices in the giant component. Indeed, the average geodesic distance, which represents the shortest path between any two vertices, is relatively short; on average, a person can

reach any other person in the giant component using eight intermediaries (or nine steps). Their tendency to form clusters is, however, relatively low, given the clustering coefficient ranges from 0 (low clustering) and 1 (high clustering), and the giant component has a coefficient of 0.12. From an STI diffusion perspective, the size of the giant component is the maximum number of people who can be reached by an STI outbreak. Whilst the commercial sex network is likely to be more connected than our data suggest, the disproportionate number of clients and sex workers in London and Milton Keynes suggests that these two cities play a relevant role in the commercial sex network and may be key for STIs prevention.

Degree assortativity refers to the correlation of the number of links that any two connected individuals have. A positive assortativity value indicates that active clients tend to meet with popular sex workers, and clients who have used the service only once or a few times tend to meet with unpopular or less popular sex workers (or sex workers with a limited number of selected clients). In our network, however, the negative assortativity score (-0.12, $p<0.001$) indicates that less active clients tend to buy services from popular (or highly reviewed) sex workers, and more active clients tend to by services from less popular sex workers.

TABLE 2 ABOUT HERE

FIGURE 2 ABOUT HERE

Vertex centrality

Table 2 compares degree, betweenness, and closeness scores of clients and sex workers. The mean degree centrality score is 1.98, i.e. individuals in the network have, on average, approximately two direct links with other network members. When we consider the role of individuals in the network, we observe that clients have, on average, more direct links than sex workers (mean degree is 2.46 and 1.66, respectively). Clients' degree centrality

scores also show wider variation than sex workers' scores (standard deviation is 5.03 and 2.09, respectively). Similar considerations apply to betweenness centrality, suggesting that clients are more central than sex workers both locally (degree) and globally (betweenness). There is less variation in clients' and sex workers' closeness scores.

Figure 2 shows the cumulative degree, betweenness, and closeness distribution for clients and sex workers. As in other sexual networks (Rocha, Liljeros, and Holme 2011; Hsieh, Kovářík, and Logan 2014), there is a small number of very active clients and popular sex workers, and a majority of actors with only one or a few connections with other individuals in the network, suggesting that immunisation strategies targeting the most central people in the network is likely to limit STI diffusion.

Vertices removal to model prevention strategies

The overall network connectivity is 0.345, i.e. 34.5% of all dyads are connected through an undirected path. Network connectivity decreases as we remove vertices randomly, based on vertices' degree, betweenness or closeness scores, or based on the role individuals have in the sex network (clients or sex workers). Of the twelve removal strategies tested, random removal strategies were the least effective (Figure 3). By removing 5% of the vertices randomly and irrespective of their role, the final network connectivity is 0.31. The random removal of clients is slightly more effective, with 28% of the dyads still connected through an undirected path after the removal of 5% of the vertices.

Interventions based on centrality scores are instead much more effective. The removal of just 2% of sex workers based on their degree or betweenness centrality leads the network connectivity scores to 0.19 and 0.16, respectively. However, strategies focusing on clients seems to have a more disruptive effect compared to those focusing on sex workers. Network connectivity drops to 0.09 (degree-based) or 0.08 (betweenness-

based) by removing 2% of the clients according to their centrality scores. Connectivity drops to 0.21 and 0.27 after closeness-based removal of clients and sex workers, respectively. Finally, type-independent strategies based on centrality scores are the most effective. After the removal of 5% of vertices with the highest betweenness centrality scores, connectedness drops to 0.009, whilst role-independent removal based on degree centrality scores leads to connectedness scores of 0.001, signalling the most effective outbreak response strategy for this online community. Connectivity after the removal of 5% of vertices with the highest closeness scores remains at 0.14.

The effectiveness of different types of vertices removal is similar when assessed using the average closeness centrality of the individuals in the network. The overall average closeness is 290. It drops to 243 after the random removal of 5% of vertices, and to 66 after closeness-based removal of the same percentage of vertices. Degree- and betweenness-based removals appear to be the most effective, with the average closeness score dropping to 2 and 9, respectively. Again, strategies focusing on clients seems to have a more disruptive effect compared to those focusing on sex workers, whether removal is based on vertices' degree, betweenness, or closeness centrality scores.

FIGURE 3 ABOUT HERE

FIGURE 4 ABOUT HERE

Discussion

We present, for the first time in the UK, a social network analysis of sex workers and clients using internet-mediated data. Similar to other studies of online sexual networks, we found a small degree of disassortativity in our network (Hsieh, Kovářík, and Logan 2014; Rocha, Liljeros, and Holme 2010). Clients that have used the service only one or a few times tend to visit highly reviewed sex workers whereas clients with many reviews tend to buy sex from less popular sex workers. In practical terms, this may suggest that

relatively new clients start out by visiting the most popular sex workers before gradually expanding their visits to less commonly reviewed ones.

We also found that the network was dominated by a giant component, which included 59% of nodes in the network and which was geographically located in London and its suburbs. A relatively high number of reviews in London is expected, but the predominant role of Milton Keynes in this online community comes as a surprise. We have identified two possible reasons for the high concentration of this online community in Milton Keynes. First, the offline commercial sex market in Milton Keynes is flourishing and very active. Milton Keynes is located 50 miles northwest of London, has a population of about 250,000 residents and is one of the fastest growing cities in the UK. It is conveniently located near several towns and cities (e.g. Northampton, Bedford, Luton, Leicester, Cambridge, and Oxford), and more than 7 million people live within an hour drive from Milton Keynes. It is an economically prosperous area characterised by the presence of many companies, and a large share of high-skilled jobs and, consequentially, high wages. All these factors contribute to Milton Keynes being a convenient place for sex workers to set up their businesses, and for clients to access such services.

Second, clients active around Milton Keynes may be unusually engaged with this online community, i.e. they tend to write more reviews than clients located elsewhere. This might be due to the intermediary role that agencies play between clients and sex workers in the area. When sex workers work through an agency, the latter takes care of the advertising, clients' screenings and bookings. If a few agencies have acquired a prominent role in Milton Keynes, they may be actively and successfully encouraging clients to write a review after the sexual encounter (Sanders et al. 2019). Irrespective of which hypothesis may be correct, the relevance of Milton Keynes in the online

community does not necessarily reflect its role in the offline sex industry. The town could be an important hub for the commercial sex industry in the UK, but online reviews might exaggerate its role. Future research, employing more traditional methods such as surveys, interviews and ethnography, might want to deepen the insight provided by our study.

The analysis also showed the role that clients and sex workers have in the network. Clients have a more central role both locally and across the network, and buying or selling sex in the two main hubs, London and Milton Keynes, is associated with high centrality scores, increasing their potential risk of contracting STIs. Counterintuitively, our analysis suggested that clients have more sexual contacts than sex workers. This last finding is likely to be artefactual given our data collection method and conservative approach in identifying sex workers. Nonetheless, the results provide an original insight into the behaviours of clients and their implication for STI transmission. Clients with high degree centrality scores, i.e. those with many sexual partners, travelled to several locations in the UK to buy sex. This may suggest that 1) prolific clients may travel looking for exclusive or potentially 'niche' services; 2) people that travel frequently for personal reasons can buy sex in different locations. Either way, prolific clients are geographically mobile and more likely than sex workers to bridge distant parts of the network (Soothill and Sanders 2005). While just a small group of clients travelled to more than two locations, some of them covered long distances. For instance, the most prolific one had 125 sexual partners across 28 different locations.

Finally, a simulation of possible preventive strategies indicated that targeting vertices using network parameters is more effective than random targeting. Interventions focusing on the most active members—whether sex workers or clients—may be thus the most effective strategy to reduce disease spread and protect the community. Notably, there was little discernible difference in whether members were targeted based on their

betweenness centrality (i.e. their role as brokers) or their degree centrality (i.e. the number of links in the network) whereas closeness-based removal underperformed compared to the other two centrality measures.

The results of this study can inform effective interventions to prevent STI transmission through sex work. They can, in fact, be used to identify individuals that are at higher risk of contracting STIs given their position in the network. While not every sexual contact leads to STI diffusion, every contact increases the probability of transmission (Zhang and Centola 2019). This probability is heightened when sexual contacts involve central (i.e. active) individuals and risky sexual practices. If central actors contract an infection, it is highly likely that many others in the network will contract it as well. The immunization of highly connected vertices can reduce the risk that others in the network will be infected. Even if an infection enters the network, it will not easily reach other people, if highly connected members resort to safer sex practices (Valente 2017; Chami et al. 2017). Whilst the effectiveness of targeting highly connected nodes over random targeting or targeting everyone is well known (e.g. Newman 2002), this study provides evidence that this may be true for online commercial sex networks, too.

Naturally, our approach presents some limitations. First, despite the continued growth in web applications for the mediation of client-sex worker relationships, our data are still likely to represent a sub-section of the entire sample of sex workers and clients engaged in off-street commercial sex in the UK, and our findings may be subject to selection bias. Whilst some network measures are likely to be relatively robust despite the bias in our data (e.g. degree distribution), others should be interpreted with caution (e.g. centrality scores for clients and sex workers).We also know very little about clients posting these reviews, and whether they correspond with the typical social media platform

user (i.e. male, young, and in managerial, administrative, or professional occupation) (Sloan et al. 2013). Second, these reviews are self-posted and completely anonymous. As with any self-report survey, there exists a possibility of fictive data. Clients can post reviews of sex workers they have never met and may not post reviews of other encounters. However, clients and sex workers can flag-up fraudulent reviews to the moderator who can, in turn, remove the reviews from the website. The number of posted reviews is used to rank clients according to their reputation and experience, which disincentives underreporting and the use of multiple accounts. Similarly, sex workers would lose their accumulated social capital and established reputation if they used multiple names for different locations and venues (Holt and Blevins 2007).[2] Third, we analysed a snapshot of the commercial sex network using reviews posted between 2014-2017 without taking time into consideration. Future research should model the dynamics of the commercial sex networks to assess how past reviews affect the creation of new links, and how the network evolves over time, with implications for STI diffusion and prevention.

Future research could also explore how network structure and composition facilitate the diffusion of behavioural norms, including safe-sex practices (Argento et al. 2016) as online communities can play a key role in sharing STI prevention messages with hard-to-reach populations (Minichiello et al. 2015). This would require an understanding of how the diffusion of safe-sex practices—which is better understood as complex contagion, i.e. contagion that requires multiple contacts and social reinforcement (Guilbeault, Becker, and Centola 2018) —can be effectively achieved in online communities of sex workers and their clients, and how multiple contagions (e.g. STI diffusion and STI prevention campaigns) would interact in the network.

---

[2] It is worth noting that, in the case of a bad early review, sex workers do have an incentive for starting a new profile.

Conclusions

The relevance of online technologies in today's sex industry makes the use of large-scale online data increasingly important to understand commercial sex networks. Online data about commercial sex contacts can provide insights into the structural and geographical characteristics of sex networks that would be otherwise prohibitive to obtain using traditional data collection methods. The results of this study show that the giant component of this online network clustered around a major conurbation. But perhaps more surprisingly, an exurb of London, Milton Keynes, played a significant role in the giant component. This suggests a role for considering possibly 'unexpected' geographical dimensions of socio-sexual networks, including possible contextual features influencing popularity of sex work services.

Our findings have several implications for public health policy and practice. Previous modelling studies of HIV prevention in sex workers have shown that small, incremental improvements in coverage of biomedical interventions (periodic condom inundation, uptake of pre-exposure prophylaxis) can effect substantial improvements in HIV incidence (Poteat et al. 2015). Indeed, under the most effective strategies simulated in this study, a low level of 'vertex removal', representing coverage of interventions to block STI transmission, yielded substantial network effects.

Future analyses should seek to understand individual, geographical, and temporal dimensions of the network as well as explore the possibility of using this community to improve population health. For example, what factors account for vertex centrality, and how did major public events (e.g. Olympic games) shape network configuration? In addition, simulation of outbreak control strategies should seek to understand the potential

impacts of structural, rather than individual, interventions on network STI diffusion. Our analysis was only able to draw on random vertex selection to simulate outbreak control effectiveness. Future analyses could draw on temporal and probabilistic models to consider more nuanced intervention strategies.

References

Argento, Elena, Putu Duff, Brittany Bingham, Jules Chapman, Paul Nguyen, Steffanie A. Strathdee, and Kate Shannon. 2016. "Social Cohesion Among Sex Workers and Client Condom Refusal in a Canadian Setting: Implications for Structural and Community-Led Interventions." *AIDS and Behavior* 20 (6): 1275–83.

Campbell, Ellsworth, and Marcel Salathé. 2013. "Complex Social Contagion Makes Networks More Vulnerable to Disease Outbreaks." *Scientific Reports* 3: 1905.

Centola, Damon. 2018. *How Behavior Spreads: The Science of Complex Contagions*. Princeton; Oxford: Princeton University Press.

Chami, Goylette F., Sebastian E. Ahnert, Narcis B. Kabatereine, and Edridah M. Tukahebwa. 2017. "Social Network Fragmentation and Community Health." *Proceedings of the National Academy of Sciences* 114 (36): E7425–31.

Crotty, Sean M., and Vanessa Bouché. 2018. "The Red-Light Network: Exploring the Locational Strategies of Illicit Massage Businesses in Houston, Texas." *Papers in Applied Geography* 4 (2): 205–27.

Cunningham, Scott, and Todd D. Kendall. 2011. "Prostitution 2.0: The Changing Face of Sex Work." *Journal of Urban Economics* 69 (3): 273–87.

Cunningham, Stewart, Teela Sanders, Jane Scoular, Rosie Campbell, Jane Pitcher, Kathleen Hill, Matt Valentine-Chase, Camille Melissa, Yigit Aydin, and Rebecca Hamer. 2018. "Behind the Screen: Commercial Sex, Digital Spaces and Working Online." *Technology in Society* 53 (May): 47–54.

Dias, Sónia. 2015. "The Role of Men Who Pay for Sex in STI Transmission: Current Knowledge and Future Directions. The Contribution of the Third UK National Survey of Sexual Attitudes and Lifestyles (Natsal-3)." *Sexually Transmitted Infections* 91 (2): 78–79.

Doherty, Irene A., Nancy S. Padian, Cameron Marlow, and Sevgi O. Aral. 2005. "Determinants and Consequences of Sexual Networks as They Affect the Spread of Sexually Transmitted Infections." *The Journal of Infectious Diseases* 191 Suppl 1 (February): S42-54.

Gezinski, Lindsay B., Sharvari Karandikar, Alexis Levitt, and Roxane Ghaffarian. 2016. "'Total Girlfriend Experience': Examining Marketplace Mythologies on Sex Tourism Websites." *Culture, Health & Sexuality* 18 (7): 785–98.

Grov, Christian, Juline Koken, Martez Smith, and Jeffrey T. Parsons. 2017. “How Do Male Sex Workers on Craigslist Differ from Those on Rentboy? A Comparison of Two Samples.” *Culture, Health & Sexuality* 19 (4): 405–21.

Guilbeault, Douglas, Joshua Becker, and Damon Centola. 2018. “Complex Contagions: A Decade in Review.” In *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, edited by Sune Lehmann and Yong-Yeol Ahn, 3–25. Computational Social Sciences. Cham: Springer International Publishing.

Handcock, Mark S, David R. Hunter, Carter T. Butts, Steven M. Goodreau, Pavel N. Krivitsky, Skye Bender-deMoll, and Martina Morris. 2005. statnet: Software Tools for the Statistical Analysis of Network Data (version 1.15 created on 01/04/2019, issued 2005. The statnet Project. www.statnet.org.

Hao, Chun, Jennifer Guida, Donald E. Morisky, and Hongjie Liu. 2015. “Family Network, Workplace Network, and Their Influence on Condom Use: A Qualitative Study Among Older Female Sex Workers in China.” *Journal of Sex Research* 52 (8): 924–35. https://doi.org/10.1080/00224499.2014.973101.

Holt, Thomas J., and Kristie R. Blevins. 2007. “Examining Sex Work from the Client’s Perspective: Assessing Johns Using on-Line Data.” *Deviant Behavior* 28 (4): 333–54.

Home Affairs Committee. 2016. “Prositution. Third Report of Session 2016–17.” London: House of Commons. Home Affairs Committee. https://publications.parliament.uk/pa/cm201617/cmselect/cmhaff/26/26.pdf.

Hsieh, Chih-Sheng, Jaromír Kovářík, and Trevon Logan. 2014. “How Central Are Clients in Sexual Networks Created by Commercial Sex?” *Scientific Reports* 4: 7540.

Jones, Kyle G., Anne M. Johnson, Kaye Wellings, Pam Sonnenberg, Nigel Field, Clare Tanton, Bob Erens, et al. 2015. “The Prevalence of, and Factors Associated with, Paying for Sex among Men Resident in Britain: Findings from the Third National Survey of Sexual Attitudes and Lifestyles (Natsal-3).” *Sexually Transmitted Infections* 91 (2): 116–23.

Klovdahl, Alden S. 2005. “Social Network Research and Human Subjects Protection: Towards More Effective Infectious Disease Control.” *Social Networks*, Ethical Dilemmas in Social Network Research, 27 (2): 119–37.

Krackhardt, David. 1994. “Graph Theoretical Dimensions of Informal Organizations.” In *Computational Organizational Theory*, edited by Kathleen M. Carley and Michael J. Prietula, 89–111. Hillsdale, NJ: Lawrence Erlbaum and Associates.

Latkin, Carl, Cui Yang, Karin Tobin, Typhanye Penniman, Jocelyn Patterson, and Pilgrim Spikes. 2011. “Differences in the Social Networks of African American Men Who Have Sex With Men Only and Those Who Have Sex With Men and Women.” *American Journal of Public Health* 101 (10): e18–23.

May, R. M., and A. L. Lloyd. 2001. “Infection Dynamics on Scale-Free Networks.” *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 64: 066112.

Mc Grath-Lone, Louise, Kimberly Marsh, Gwenda Hughes, and Helen Ward. 2014. “The Sexual Health of Female Sex Workers Compared with Other Women in England: Analysis of Cross-Sectional Data from Genitourinary Medicine Clinics.” *Sexually Transmitted Infections* 90 (4): 344–50.

Newman, M. E. J. 2002. “The Spread of Epidemic Disease on Networks.” *Physical Review E* 66 (1): 016128. https://doi.org/10.1103/PhysRevE.66.016128.

———. 2003. “Mixing Patterns in Network.” *Physical Review E: Statistical, Non-Linear, and Soft Matter Physics* 67 (2): 1–13.

Noack-Lundberg, Kyja, Pranee Liamputtong, Brahm Marjadi, Jane Ussher, Janette Perz, Virginia Schmied, Tinashe Dune, and Eloise Brook. 2019. “Sexual Violence and Safety: The Narratives of Transwomen in Online Forums.” *Culture, Health & Sexuality* 0 (0): 1–14.

Opsahl, Tore. 2009. “Structure and Evolution of Weighted Networks.” Ph.D., Queen Mary, University of London. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.507253.

———. 2011. “Triadic Closure in Two-Mode Networks: Redefining the Global and Local Clustering Coefficients.” *ArXiv:1006.0887 [Physics]*, May. http://arxiv.org/abs/1006.0887.

Platt, Lucy, Pippa Grenfell, Chris Bonell, Sarah Creighton, Kaye Wellings, John Parry, and Tim Rhodes. 2011. “Risk of Sexually Transmitted Infections and Violence among Indoor-Working Female Sex Workers in London: The Effect of Migration from Eastern Europe.” *Sexually Transmitted Infections* 87 (5): 377–84.

Poteat, Tonia, Andrea L Wirtz, Anita Radix, Annick Borquez, Alfonso Silva-Santisteban, Madeline B Deutsch, Sharful Islam Khan, Sam Winter, and Don Operario. 2015. “HIV Risk and Preventive Interventions in Transgender Women Sex Workers.” *The Lancet* 385 (9964): 274–86.

R Core Team. 2019. R: A Language and Environment for Statistical Computing (version 3.6.1)., issued 2019. Vienna: R Foundation for Statistical Computing. https://www.R-project.org.

Rocha, Luis E. C., Fredrik Liljeros, and Petter Holme. 2011. “Simulated Epidemics in an Empirical Spatiotemporal Network of 50,185 Sexual Contacts.” *PLoS Computational Biology* 7 (3): e1001109.

Rocha, Luis E.C., Fredrik Liljeros, and Petter Holme. 2010. “Information Dynamics Shape the Sexual Networks of Internet-Mediated Prostitution.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (13): 5706–11.

Rothenberg, Richard, Tina Dan My Hoang, Stephen Q. Muth, and Richard Crosby. 2007. “The Atlanta Urban Adolescent Network Study: A Network View of STD Prevalence.” *Sexually Transmitted Diseases* 34 (8): 525–31.

Sanders, T., J. Scoular, J. Pitcher, R. Campbell, and S. Cunningham. 2019. “Beyond the Gaze and Well Beyond Wolfenden and: The Practices and Rationalities of Regulating/ Policing Sex Work in the Digital Age.”

Sanders, Teela, Jane Scoular, Rosie Campbell, Jane Pitcher, and Scott Cunningham. 2018. *Internet Sex Work*. https://www.palgrave.com/gp/book/9783319656298.

Schrager, Sheree M., Carl A. Latkin, George Weiss, Katrina Kubicek, and Michele D. Kipke. 2013. “High-Risk Sexual Activity in the House and Ball Community: Influence of Social Networks.” *American Journal of Public Health* 104 (2): 326–31.

Shannon, Kate, Shira M. Goldenberg, Kathleen N. Deering, and Steffanie A. Strathdee. 2014. “HIV Infection among Female Sex Workers in Concentrated and High Prevalence Epidemics: Why a Structural Determinants Framework Is Needed.” *Current Opinion in HIV and AIDS* 9 (2): 174–82.

Shushtari, Zahra Jorjoran, Seyed Ali Hosseini, Homeira Sajjadi, Yahya Salimi, Carl Latkin, and Tom A. B. Snijders. 2018. “Social Network and HIV Risk Behaviors in Female Sex Workers: A Systematic Review.” *BMC Public Health* 18 (1): 1020.

Sloan, Luke, Jeffrey Morgan, William Housley, Matthew Williams, Adam Edwards, Pete Burnap, and Omer Rana. 2013. "Knowing the Tweeters: Deriving Sociologically Relevant Demographics from Twitter." *Sociological Research Online* 18 (3): 7.
Soothill, Keith, and Teela Sanders. 2005. "The Geographical Mobility, Preferences and Pleasures of Prolific Punters: A Demonstration Study of the Activities of Prostitutes' Clients:" *Sociological Research Online*.
Valente, Thomas. 2017. "Putting the Network in Network Interventions." *Proceedings of the National Academy of Sciences* 114 (August): 201712473.
Valente, Thomas W., and Stephanie R. Pitts. 2017. "An Appraisal of Social Network Theory and Analysis as Applied to Public Health: Challenges and Opportunities." *Annual Review of Public Health* 38 (1): 103–18.
Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York, NY: Cambridge University Press.
Youm, Yoosik. 2015. "Sexually Transmitted Infections: Social Network Analysis." *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, March, 853–61.
Zhang, Jingwen, and Damon Centola. 2019. "Social Networks and Health: New Developments in Diffusion, Online and Offline." *Annual Review of Sociology* 45 (1):91–109

Acknowledgments

This work was previously presented at the XXXVIII Sunbelt Conference, June 26-July 1, 2018, Utrecht, The Netherlands. We would like to thank Andrew Gillard for his assistance in the data collection.

Disclosure statement

We have no conflicts of interest to declare.

Funding

This research was partially supported by Cardiff University School of Social Sciences Research Committee funding.

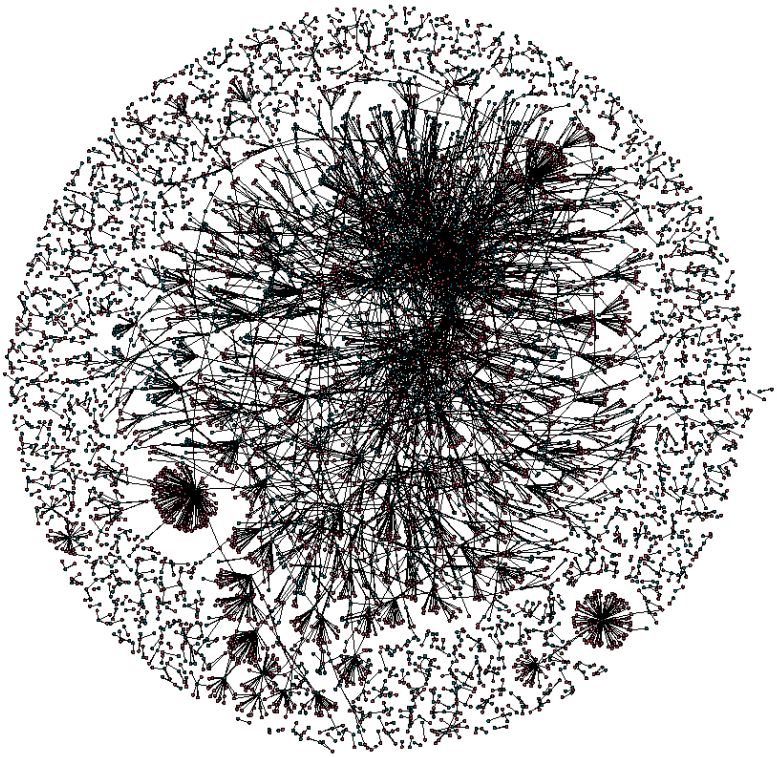
Ethics approval

This research has been approved by the School of Social Sciences Research Ethics

Committee of Cardiff University (REF: SREC/2745).

Figures

Figure 1. Sociogram of the commercial sex network



Note: blue vertices represent clients whilst red vertices represent sex workers.

Figure 2. Degree, betweenness, and closeness distribution for sex workers and clients

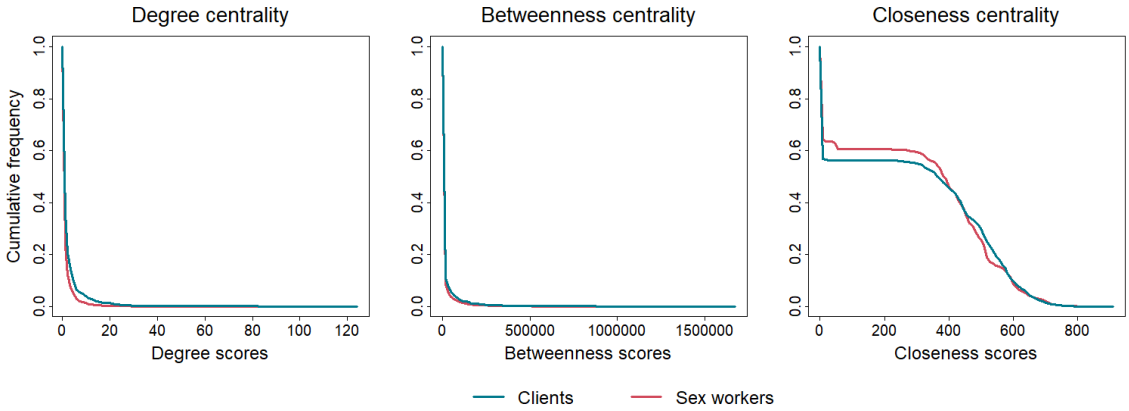


Figure 3. Impact of different strategies on the level of connectivity of the network

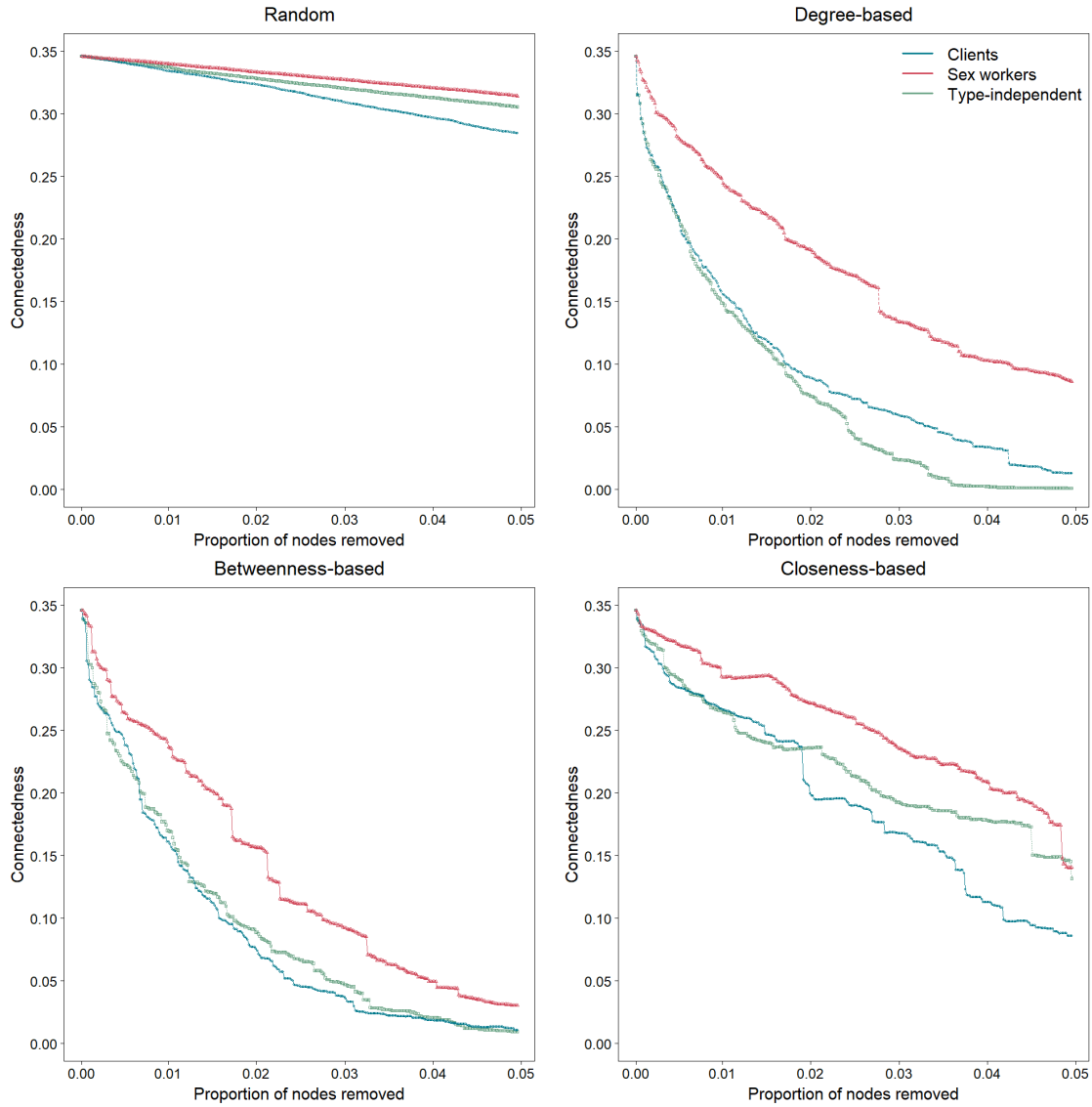
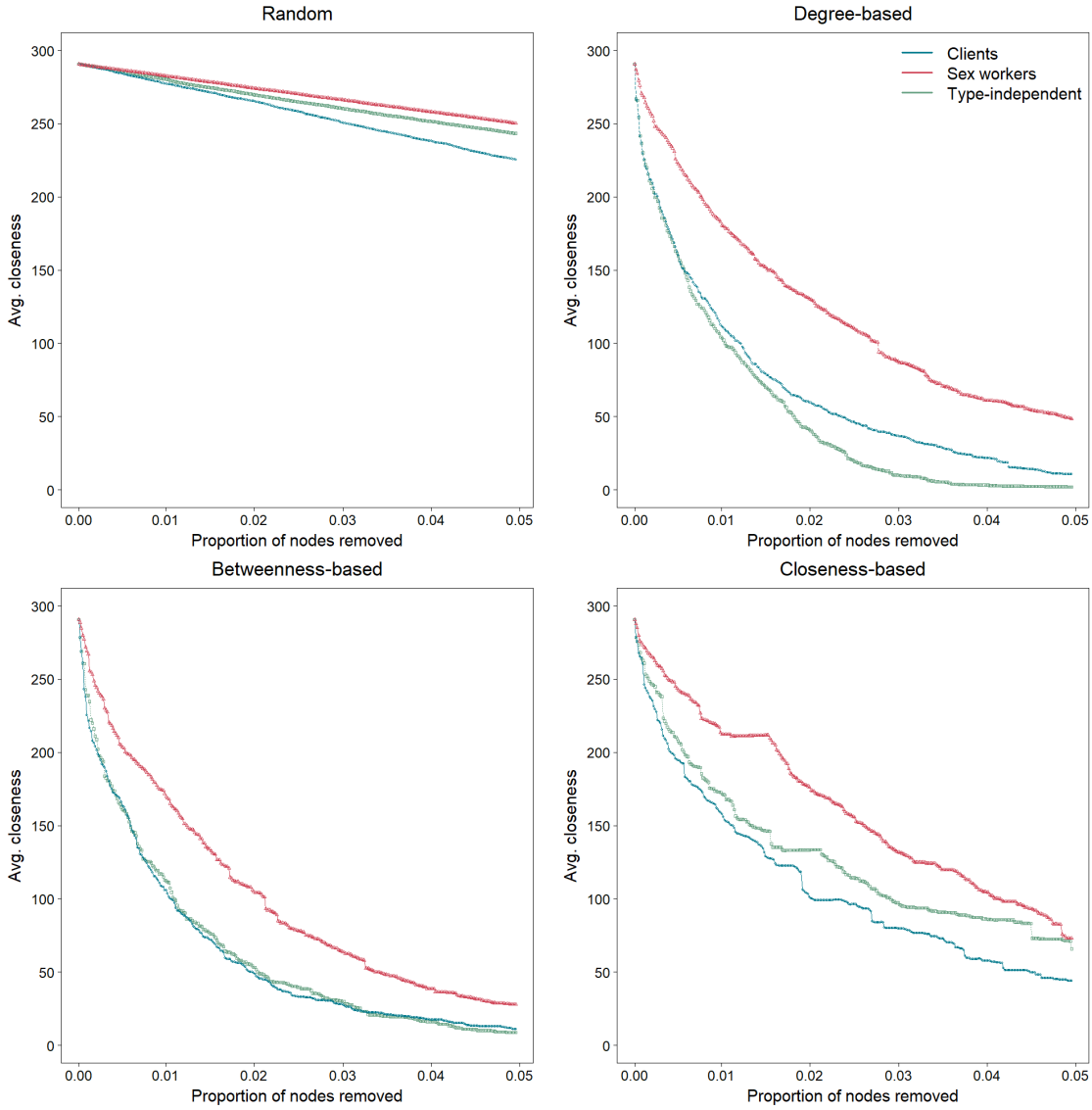


Figure 4. Impact of different strategies on the average closeness degree of the network



Tables

Table 1. Descriptive statistics of the sex network

Number of vertices	6477
Sex workers	3870 (60%)
Clients	2607 (40%)
Number of venues	1656
Locations	328
Number of edges	6426
Assortativity	-0.12 ($p<0.001$)
Giant component – vertices	3807 (59%)
Giant component – edges	4715 (73%)
Average geodesic distance (in GC)	9
Diameter (in GC)	29
Clustering (in GC)	0.12
Second largest component – vertices	125

Table 2. Clients and sex workers position in the network

	Degree			**Betweenness**			**Closeness**		
	TOT	**SW**	**CL**	**TOT**	**SW**	**CL**	**TOT**	**SW**	**CL**
Mean	1.98	1.66	2.46	8785	7135	11234	291	296	283
St. Dev	3.60	2.09	5.03	44629	32964	57666	253	248	260
Median	1	1	1	0	0	0	379	386	366
Min	1	1	1	0	0	0	1	1	1
Max	124	33	124	1671442	733232	1671442	909	823	909
Skewness	14.30	6.91	12.03	16.17	10.65	15.59	-0.06	-0.11	0.02

Figures caption

Figure 1. Sociogram of the commercial sex network

Figure 2. Degree, betweenness, and closeness distribution for sex workers and clients

Figure 3. Impact of different strategies on the level of connectivity of the network

Figure 4. Impact of different strategies on the average closeness degree of the network