PLOS ONE

On limitations of uniplex networks for modeling multiplex contagion

Nicholas W. Landry [ID]1,2,3, **jimi adams** [ID]4,5 *

1 Vermont Complex Systems Center, University of Vermont, Burlington, VT, United States of America, **2** Department of Mathematics and Statistics, University of Vermont, Burlington, VT, United States of America, **3** Department of Applied Mathematics, University of Colorado Boulder, Boulder, CO, United States of America, **4** Department of Health & Behavioral Sciences, University of Colorado Denver, Denver, CO, United States of America, **5** Institute of Behavioral Science, University of Colorado Boulder, Boulder, CO, United States of America

* jimi.adams@ucdenver.edu

Abstract

Many network contagion processes are inherently multiplex in nature, yet are often reduced to processes on uniplex networks in analytic practice. We therefore examine how data modeling choices can affect the predictions of contagion processes. We demonstrate that multiplex contagion processes are not simply the union of contagion processes over their constituent uniplex networks. We use multiplex network data from two different contexts—(1) a behavioral network to represent their potential for infectious disease transmission using a "simple" epidemiological model, and (2) users from online social network sites to represent their potential for information spread using a threshold-based "complex" contagion process. Our results show that contagion on multiplex data is not captured accurately in models developed from the uniplex networks even when they are combined, and that the nature of the differences between the (combined) uniplex and multiplex results depends on the specific spreading process over these networks.

Introduction

Studies of contagion on networks regularly rely on data that only represent one type of relationship at a time. These sorts of data can be helpful for understanding contagion dynamics on relationships of that type. However, when the underlying contagion processes are multiplex, they are less well suited for understanding the ultimate extent of spread over a population and the timescale over which these processes occur.

As an example, consider research aiming to model an epidemic of a particular sexually transmitted infection such as human immunodeficiency virus (HIV). HIV is transmissible via sexual contact, shared needles, and other exposures to blood or other bodily fluids. In network terms, this means that any spread of HIV through the population requires an unbroken chain of susceptible cases exposed to those who are currently infectious [1]. That chain of exposure could be contained within relationships of a particular type (e.g., could be completely

composed of sexual contacts), or could combine relationships of multiple types (i.e., sexual and needle-sharing contacts).

In many populations, members have active contacts of both types, and they are not directly overlapping—for example, not all needle-sharing contacts are with the same partners involving sexual contacts [2]. In that case, studying only one type of relationship at a time could substantially distort any predictions of epidemic extent within the population. As a simple example, suppose person 1 in Fig 1 is currently infected, and we want to know the likelihood of person 4 becoming infected. If we examine only sexual relationships (denoted by solid lines), both 1 and 4 are sexually active with a single partner (2 and 3, respectively) and those direct relationships would be accurately represented with data on the sexual network along with any models using those data. However, the potential infection-carrying relationship between 2 and 3 providing an indirect route of transmission between 1 and 4 would not be observed, because it stems from a needle-sharing relationship (dashed line). With only data on sexual relationships, we would be underestimating 1's likelihood of *indirectly* infecting 4, and even misunderstanding how these nodes' sexual partnerships contribute to that possible route of transmission. We see that even in this simple example, only considering one tie type can lead to very different predictions of the likelihood of infection and the resulting epidemic extent.

Similarly, there has been a recent proliferation of studies examining propagation of memes, ideas, even voting behavior across social media sites like Facebook or Twitter [3–5]. However, while people mix their social media usage and have non-overlapping connectivity patterns across platforms [6], these studies overwhelmingly focus on modeling contagion over a single platform at a time. It is reasonable to expect this approach could potentially give a limited picture of the complete contagion potential across the population, as in the sex/needle example for STIs above. Moreover, simply increasing the coverage of the sample used will not overcome these potential sampling biases [7]. For example, people may primarily use one platform with their friends and family (ties between {1, 2} and {3, 4} in Fig 1), while using another with their professional contacts (tie {2, 3} in Fig 1). In that case, studies that rely on detailed data from a single platform at a time would miss important cross-platform linkages that are vital to understanding the contagion dynamics across the population of interest. Information and misinformation have been shown to spread between different social media platforms, reinforcing the importance of understanding limitations in modeling multiplex contagion [8].

In this paper, we demonstrate the potential limitations of inference about population contagion potential and timescale from networks of a single relationship type (uniplex networks), when contagion processes realistically involve spreading over relationships of multiple types (multiplex networks).

We proceed as follows: first, we motivate the importance of these limitations on contagion modeling; second, we discuss the data sets and contagion processes considered; third, we discuss our approach in quantifying discrepancies in the epidemic trajectories for different choices of data; fourth, we present numerical results validating our premise; and lastly, we discuss limitations and interpretations of this study.

Fig 1. A simple multiplex network. The line style corresponds with the relationship type. The node color represents its current status: infected (red), susceptible (white), and unknown (black).

https://doi.org/10.1371/journal.pone.0279345.g001

Motivation

Our central aim is to show that researchers' ability to model contagion is strongly constrained by choices in the data collection and representation process. Many studies consider a single risk behavior or platform for potential disease/idea spread. This choice can result from simple limitations of data availability, or may draw on the recognition that domain expertise often improves our capacity to collect high quality data on particular topics (e.g., drug sharing networks have considerable barriers to data elicitation [9]). It is important to understand the limitations introduced by these choices when predicting the extent to which behaviors or diseases will spread. We draw on two simple case studies to explore the discrepancies between the contagion extent resulting from spreading over a single relationship type compared to a model with two different relationship types.

Many studies have presented theoretical approaches for describing contagion dynamics over multiplex networks [10–13]. These studies have been crucial in understanding of dynamics on multiplex networks, but take a fundamentally different approach compared to ours. This study focuses not on predicting the behavior of a multiplex network, but rather exploring the effect of the underlying data representation on the limitations that arise when modeling contagion processes from data representing a single relationship type. This is similar to the approach in Ref. [14] where the authors examine how well networks constructed from different data types effectively capture transmission potential. We show that one cannot generate accurate expectations of epidemic extent from uniplex data, even if combining results across layers to reconstruct the multiplex reality, as an estimate of their potential combined effects—and those discrepancies vary by the nature of the contagion process.

Data

We use two data sets to represent multiplex networks. For each, we examine (1) the entire multiplex network and (2) the multiplex network constructed with nodes that are members of the largest connected component of the network when considering both relationship types.

The first data set comes from "Project 90" in Colorado Springs, which studied a network of commercial sex workers, people who inject drugs, and their sex and needle-sharing partners between 1988–1992 [15]. Here we use the complete network of 7,677 individuals representing all respondents and the 'risk behavior' partners they nominated as well as the largest connected component of 4,385 individuals. Although these data were collected over five different waves, we aggregated them into a single network for our purposes. The other data set represents a multiplex network of online social network site interactions among a sample of 1,672 Twitter and Foursquare users [16] hereafter, 'JOAAP' as well as the largest connected component of 1,564 individuals. These data sets are available for download from Refs. [17, 18], respectively. Summary network properties are presented in Table 1.

In the Project 90 data, we considered sexual ties as one type of tie and drug or needle-sharing ties as a second tie type. When we refer to the *multiplexed data*, we aggregate these uniplex layers by forming a link between two individuals if they have a sexual relationship or a drug or needle-sharing relationship, whereas the uniplex networks are formed from each tie type separately.

The JOAPP data, which are much higher density, representing Foursquare co-check-ins (frequenting the same locales) as one type of tie and "follower" links on Twitter as a second type of tie. We symmetrized the Twitter data for consistent analysis, even though these data are inherently directional.

Analyses begin with the observed multiplex Project 90 and JOAAP data sets [15, 18]. We then decompose each of these multiplex networks into their respective uniplex networks with

Table 1. Statistics of the chosen networks.

Data set	N	$\langle k \rangle$	$\langle c \rangle$	C	κ
Full data set					
1. Project 90	7677	1.70	2495.93	1.59×10^{-2}	2.13×10^{-2}
2a. Project 90—Sex	7677	0.85	466.44	7.78×10^{-4}	4.26×10^{-2}
2b. Project 90—Drugs	7677	1.17	1039.01	1.53×10^{-2}	3.08×10^{-2}
3. JOAPP	1672	45.7	1672	0.519	3.32×10^{-2}
4a. JOAPP—Foursquare	1672	35.22	1672	0.613	4.32×10^{-2}
4b. JOAPP—Twitter	1672	16.85	1455.58	0.123	8.99×10^{-2}
Largest component of multiplexed data					
1. Project 90	4375	2.46	4375	2.62×10^{-2}	1.71×10^{-2}
2a. Project 90—Sex	4375	1.22	816.69	1.37×10^{-3}	3.46×10^{-2}
2b. Project 90—Drugs	4375	1.71	819.74	2.52×10^{-2}	2.46×10^{-2}
3. JOAPP	1564	36.21	1564	0.302	2.66×10^{-2}
4a. JOAPP—Foursquare	1564	24.67	1538.12	0.317	3.91×10^{-2}
4b. JOAPP—Twitter	1564	18.02	1556.02	0.131	5.35×10^{-2}

each tie-type representing a unique network. For each data set, this produces three networks. For the Project 90 data set, this produces (1) the multiplex composite risk behavior network, (2a) the sexual contact uniplex network, and (2b) the needle-sharing uniplex network. For the JOAAP data set, this produces (3) the multiplex composite online network, (4a) the Foursquare uniplex network, and (4b) the Twitter uniplex network. All analyses described below are then conducted on each of these networks separately, and the results compared between the multiplex networks (1 and 3) and the union of the uniplex layers (2a and 2b, and 4a and 4b, respectively).

For each data set, we extract the list of node labels by tabulating all the nodes that are connected to either link type. For each type of network (uniplex or multiplex), we start with this node list and then add the links corresponding to the desired tie types. Although in principle one could weight the links according to the frequency of their occurrence, for simplicity, we consider unweighted uniplex and multiplex networks, even if two ties of different types connect the same node pairs, accounting for why the average degree of the two uniplex networks do not sum to the average degree of the multiplexed network.

In our study, we use the Project 90 data as a model of a behavioral network over which epidemics can spread and the JOAPP data as a model of online networks for behavior adoption. Table 1 presents descriptive statistics for each data set. In Table 1, N denotes the number of nodes in the network and $\langle k \rangle$ denotes the average degree. In the susceptible—infected contagion model, the size of the connected component of which a seed node is a member determines the epidemic size. For this reason, we compute the average size of a connected component given a seed node selected uniformly at random with the expression $\langle c \rangle = \sum_{i}^{N_c} c_i^2 / N$, given connected components $c_1, \ldots, c_{N_c}$. Likewise, for complex contagion, the neighborhood of a randomly selected node can be predictive of how behavioral contagion will spread and we compute the average clustering coefficient, C, as defined by the authors in Ref. [19]. It can be helpful to measure the extent of the overlap between two layers in a multiplex network and there are a wide variety of ways of computing this statistic [20]. In Table 1, we compute a slight modification of the *degree of multiplexity* defined in Ref. [21] and measure the fraction of links that exist in both layers with respect to the number of links in each of the uniplex networks as well as the multiplexed network. We denote this quantity as κ.

Contagion processes

In this study, we considered two models of contagion: a susceptible—infected (SI) model representing a epidemiological simple contagion process, and a threshold model representing a complex contagion process, which represent two common model forms of contagion processes.

For the SI model, we define two states: susceptible (S) and infected (I). A susceptible node can be infected by one of its infected neighbors with probability $\beta\Delta t$, where β is the infection rate and Δt is the interval at which the node states are updated [22]. In this study, $\Delta t = 1$ week and we update the nodes at the next time synchronously. Once a susceptible node is infected, it will remain infected. Thus, for a connected network, given enough time, every node will eventually become infected.

The threshold model (also known as "complex contagion") is a common model for behavior adoption on networks [23, 24]. For the threshold model, we define, as above, two states: non-adopting (S) and adopting (I). We fix an adoption threshold τ between 0 and 1. A non-adopting node adopts the opinion if the fraction of its neighbors who have already adopted is larger than τ. Once a node has adopted it will not change its state to become non-adopting. We update the opinions of all the nodes at the next time step synchronously; the threshold model is deterministic once the initial state and threshold value have been specified. Unlike the SI model, a contagion spreading via the threshold process may not reach the entirety of a network, even if it is connected [25].

For both models, we select seed nodes uniformly at random to infect and this defines the initial state of our system. We run these contagion processes for a sufficient duration to ensure the epidemic extents have reached equilibrium.

Approach

We demonstrate how network data representation choices alter the epidemic extent and rate of contagion spread. The data sets described above are inherently multiplex in nature and serve as case studies in answering our central question.

We start by constructing networks from the data sets as described above. For each data set, we produce the observed multiplex network containing both tie types, and two constituent uniplex networks from each data set.

For each simulation run, we fix a single seed node for each of the four settings of interest. We begin by running these respective contagion processes (the SI model on the Project 90 data set and the threshold model on the JOAPP data set) for the following: (1) the respective multiplex network data, (2) and (3) each of the constituent uniplex layers decomposed from those multiplex networks (as described above), and (4) the union of the infected nodes in each layer of (2) and (3) for each time step in the simulated contagion processes. We present results for the proportion of infected network members at each time step and final epidemic extent. For each setting combination, we generate many realizations of these simulations to form an ensemble of time series.

Results

For all contagion processes and each parameter value, we set the time step to a week, i.e., $\Delta t = 1$ and infected a single node at random initially (although we present additional results in the Supporting Information using different numbers of seed nodes). We ran 1000 simulations to form an ensemble for robustness.

For the SI process, we used the following infection rates (which, because $\Delta t = 1$, are also infection probabilities): $\beta = 1/75, 1/50, 1/30, 1/20$ (simulations for wider parameter ranges in
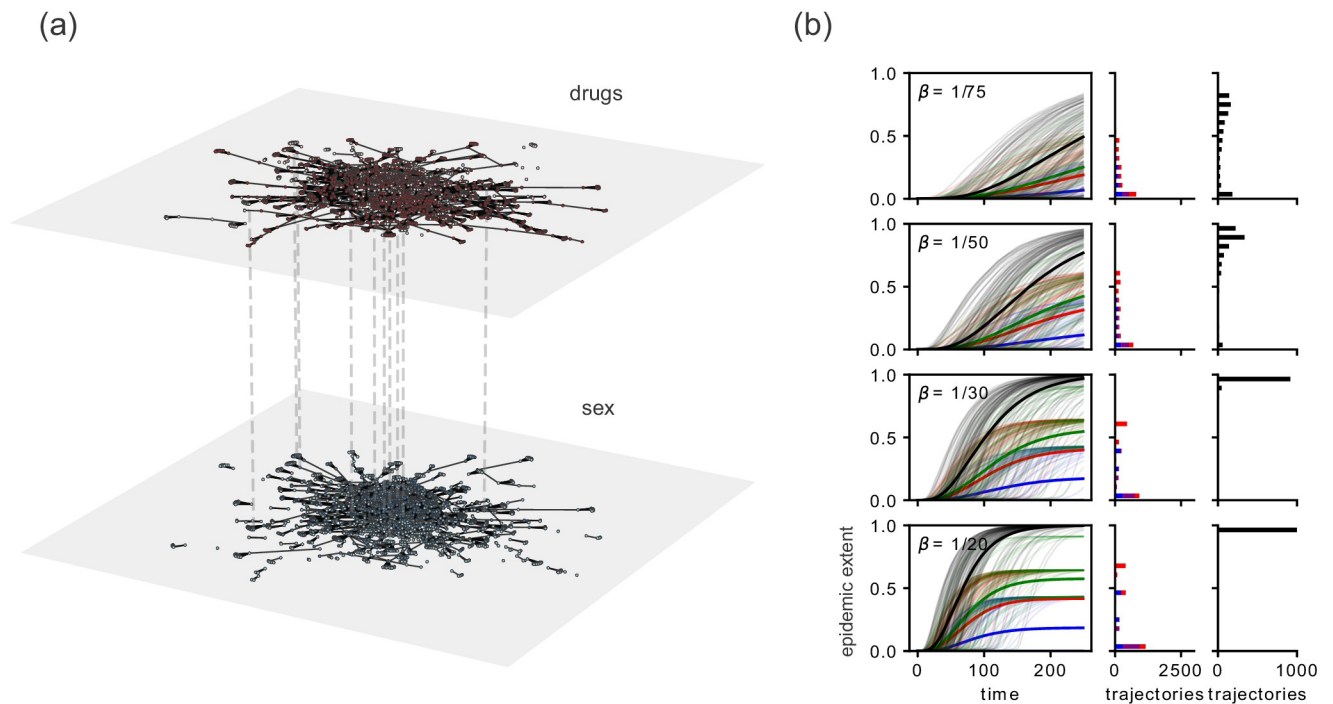
(a)

(b)

Fig 2. Comparing the epidemic extent for different choices of network data for the Project 90 data set. (a) A network visualization of the Project 90 data set. The dashed lines are not exhaustive, but illustrate that the nodes in each layer are identical. The lightly colored nodes in each layer are disconnected from the largest connected component in that layer. One can see that each layer in this multiplex network has different contact structure. (b) The SI model simulated on the Project 90 data. The left panel displays a plot of the epidemic extent with respect to time for the uniplex sex network (blue), uniplex drug and needle network (red), the union of uniplex networks (green), and the multiplexed data (black). The thick lines indicate the average epidemic extent and the thin lines are a random sample of 100 individual realizations of this model. The center column illustrates the histogram of epidemic extents for the number of nodes solely accessed from the sex network (blue), the drug network (red), and both (magenta). The right column illustrates the distribution of epidemic extents for the multiplexed data.

https://doi.org/10.1371/journal.pone.0279345.g002

the Supporting Information provide a robustness check). We simulated up until a time of t_{max} = 250 (i.e., 5 years; although we present longer simulations in the Supporting Information as a robustness check). For the threshold process, we used the following adoption thresholds: τ = 1/8, 1/10, 1/12, 1/15 (simulations for wider parameter ranges in the Supporting Information provide a robustness check). We simulated until t_{max} = 30 (although we present longer simulations in the Supporting Information as a robustness check).

To examine our central question, we draw the reader's attention to comparisons between the black and green curves in Figs 2(b) and 3(b). For the SI model, Fig 2(b) shows that the multiplex data leads to a larger epidemic extent than when we consider the union of the separate uniplex processes. This result demonstrates that to capture the "true" potential of the contagion process on the multiplex network, it is not enough to simply combine the results from their constituent uniplex layers. As shown in Fig 1, simply considering relationships of one type ignores how contagion can rely on the complementarities across relationships of different types. These uniplex networks, when combined into the multiplex nature of reality, unlock connection patterns that reach more of the population than can be accessed by the combination of the two independently. This indicates, as illustrated in Fig 1, that there is "leap-frog" behavior occurring, where a contagion must pass through connections of alternating types to reach certain nodes.

Looking at the histogram of final extents, we see that for the multiplexed data, by construction, the contagion will spread to the entire population given enough time. The uniplex data
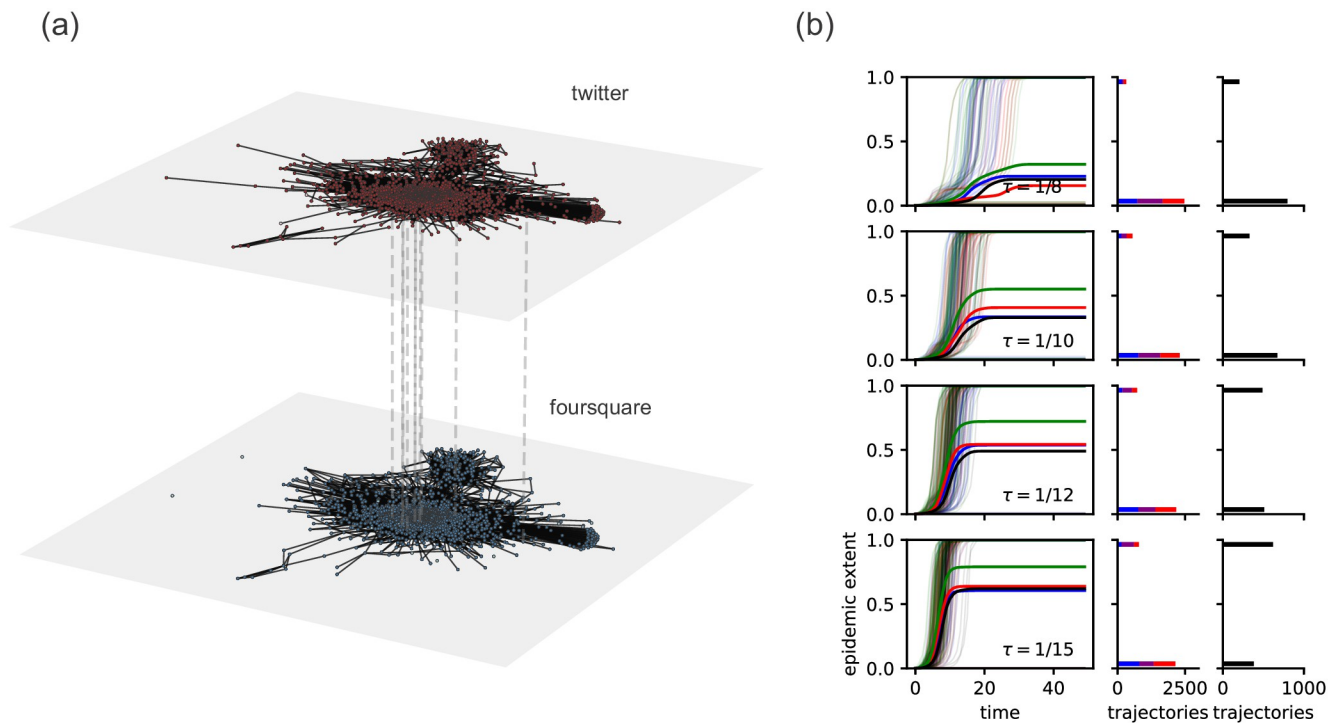
(a)   (b)

Fig 3. Comparing the epidemic extent for different choices of network data for the JOAPP data set. (a) Network visualization of the JOAPP data set. The dashed lines are not exhaustive, but illustrate that the nodes in each layer are identical. The light colored nodes in each layer are disconnected from the largest connected component in that layer. (b) The threshold model on the JOAPP data. As in Fig 2(b) the left column displays a plot of the epidemic extent with respect to time for the uniplex Twitter network (blue), uniplex Foursquare network (red), the union of uniplex networks (green), and the multiplexed data (black). The thick lines indicate the average epidemic extent and the thin lines are a random sample of 100 individual realizations of this model. The center column illustrates the histogram of epidemic extents for the number of nodes solely accessed from the Twitter network (blue), the Foursquare network (red), and both (magenta). The right column illustrates the distribution of epidemic extents for the multiplexed data.

sets do not reach the entirety of the population indicating that the uniplex networks have components that are smaller than those of the full multiplex networks.

For the threshold model, the situation is a bit less straightforward. In particular, here the results are more bimodal in that the ultimate contagion extent typically either (a) stalls before taking off at all or (b) reaches the vast majority of the population. This result reflects that both the connectedness and the density of the network determine the epidemic extent (see for example, Ref. [25]). This result again demonstrates that understanding contagion processes on uniplex data representation, even when combining them, does not straightforwardly translate to account for contagion on the related multiplex network. Indeed, in Fig 3(b), we see that for each parameter value the epidemic extent for the multiplexed data may be lower than that of both the uniplex network ($\tau = 1/10, 1/12$), larger than that of the Twitter uniplex network but smaller than that of the Foursquare uniplex network ($\tau = 1/15$), and larger than that of the Foursquare uniplex network but smaller than that of the Twitter uniplex network ($\tau = 1/15$). This is likely a complicated interplay between the connectedness and density of these networks. On the one hand, the uniplex networks are more disconnected than the multiplexed network. On the other hand, the multiplexed network is denser than each of the uniplex networks, which discourages the spread of contagion for a given threshold value. If we set $\tau = 0$, this removes the first of these two factors and the relations between the epidemic extents becomes the same as for the SI model. By definition, the epidemic extent for the union of these is larger and we see that considering both types of relationships separately

may be a *worse* estimate than simply considering a single uniplex network (for example $\tau = 1/10$, $1/12$ in Fig 3(b)).

Discussion

In this paper, we have drawn on two cases to show how modeling inherently multiplex contagion processes with uniplex network representations can misrepresent the predicted fraction of the population that a contagion process will reach. We illustrate this limitation separately for models of epidemiological and behavioral contagion to highlight the effect of the data representation on the resulting epidemic behavior. For simplicity, within each model we used the same transmission parameters for each link type. Contagion on empirical networks may have different rates of spread depending on the type of link through which contagion is transmitted (e.g., needle sharing transmits most STIs more efficiently than sexual contacts). In addition, contagion spread can be modeled by different processes spreading on each layer. For example, one might model the spread of awareness as a behavioral process on one network layer and the spread of disease on another network layer [26].

In the case of the SI model, the multiplex epidemic extent was consistently larger than that of the union of the two constituent uniplex layers. The magnitude of this discrepancy depends on the extent of overlap in the connected components from each uniplex layer. The case where each uniplex network has the same connected components will lead to the same epidemic extent. This is a trivial case where one link type may be a proxy for another link type, which may limit the usefulness of a multiplex representation, although is likely to be uncommon empirically. For the threshold process, gaps remain, but differ in their nature. Whether the union of the constituent uniplex data underestimates or overestimates the true contagion extent depends upon the threshold levels, and these results depend on combinations of network structural characteristics in ways that are important to examine further.

It may be fruitful to estimate how the magnitude of the under or overestimates seen here differ when examining data sets with different structural features. To this end, modeling multiplex data sets using random network models may be helpful in predicting the quantitative differences in the epidemic extent for different data modeling choices.

Our results offer evidence supporting our premise that uniplex data is inadequate for modeling inherently multiplex processes. We offer this study as a cautionary tale for researchers modeling the spread of contagion: the choice of network data is an important assumption baked into contagion models and should be carefully considered. As a more positive recommendation, a primary takeaway of our results is that future data collection efforts should prioritize faithfully capturing the multiplex realities of the underlying processes intended to be examined as in Ref. [14], rather than relying solely on data of a single type.

Supporting information

S1 Fig. Full data sets. We show that our results hold for the full (a) Project 90 and (b) JOAPP data sets, not just the largest component of the multiplexed data. In contrast to Figs 2b and 3b in the main text, the multiplexed data set is no longer fully connected, leading to some epidemic trajectories that reach very few nodes, resulting in the bimodal distribution of epidemic extents. For the Project 90 data set, we see in S1(a) Fig that the relative epidemic extents are preserved when compared with the epidemic extents of the largest connected component of the multiplexed data. This is not the case with the JOAPP data set in S1(b) Fig, but as discussed in the main text, this is to be expected due to the two competing factors, network density and connectedness, that determine the epidemic extent for the threshold contagion process. For

more details on these plots, see Figs 2b and 3b in the main text.
(TIF)

S2 Fig. Larger parameter ranges. We simulate each contagion process for a wider range of parameters for (a) the Project 90 data set and (b) the JOAPP data set to verify that our results are not dependent on the choice of parameters. For the SI model on the Project 90 data set in S2(a) Fig, we notice the same differences in epidemic extents, albeit on different time scales. This should be anticipated; if we plot the epidemic extent with respect to βt, the average of each time series should be the same. For the threshold process on the JOAPP data set in S2(b) Fig, there are parameter values where the epidemic extents are trivially the same. First, if we choose a threshold greater than the maximum possible fraction of infected neighbors that the network structure and number of seed nodes allow, then contagion will never occur. Second, if the threshold is low enough and the multiplex and uniplex representations are all fully connected, then the entire population will be infected no matter the data representation. For more details on these plots, see Figs 2b and 3b in the main text.
(TIF)

S3 Fig. Different number of seeds for the Project 90 data set. We show that our results hold for (a) 2, (b) 5, and (c) 10 seeds as well as a single seed node as presented in the main text. In S3 Fig, we see that the relative differences in the epidemic extents are preserved. The largest difference is that with a larger number of seed nodes, the likelihood that there is an epidemic trajectory that spreads to very few nodes is much smaller as can be seen in the figure. For more details on these plots, see Figs 2b and 3b in the main text.
(TIF)

S4 Fig. Different number of seeds for the JOAPP data set. We show that our results hold for (a) 2, (b) 5, and (c) 10 seeds as well as a single seed node as presented in the main text. In S4 Fig, we see that, as in S3 Fig, there is a smaller chance of trajectories dying out. In addition, we see that increasing the number of seed nodes effectively raises the maximum threshold for which the contagion will spread to the entire network (for example, $\tau = 1/10$ in S4 Fig) and lead to trivial results as discussed prior. For more details on these plots, see Figs 2b and 3b in the main text.
(TIF)

S5 Fig. Full temporal extent. We run the simulations for a long enough time to remove any temporal censoring for (a) smaller values of β for the SI model on the Project 90 data set and for (b) larger values of τ for the threshold model on the JOAPP data set. In S5 Fig, we see that the epidemic extents are consistent with our results in the main text. For the SI model, this should be expected as explained prior; rescaling time by the infection probability should yield very similar epidemic responses in expectation. For more details on these plots, see Figs 2b and 3b in the main text.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: jimi adams.

Data curation: Nicholas W. Landry.

Formal analysis: Nicholas W. Landry.

Methodology: Nicholas W. Landry, jimi adams.

Supervision: jimi adams.

Visualization: Nicholas W. Landry.

Writing – original draft: Nicholas W. Landry, jimi adams.

Writing – review & editing: Nicholas W. Landry, jimi adams.

References

1. Moody J, Adams J, Morris M. Epidemic Potential by Sexual Activity Distributions. Network Science. 2017; 5(4):461–475. https://doi.org/10.1017/nws.2017.3 PMID: 29449942

2. adams j, Moody J, Morris M. Sex, Drugs, and Race: How Behaviors Differentially Contribute to the Sexually Transmitted Infection Risk Network Structure. American Journal of Public Health. 2013; 103 (2):322–329. https://doi.org/10.2105/AJPH.2012.300908 PMID: 23237162

3. Aral S, Muchnik L, Sundararajan A. Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks. Proceedings of the National Academy of Sciences. 2009; 106 (51):21544–21549. https://doi.org/10.1073/pnas.0908800106 PMID: 20007780

4. Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, Settle JE, et al. A 61-Million-Person Experiment in Social Influence and Political Mobilization. Nature. 2012; 489(7415):295–298. https://doi.org/10.1038/nature11421 PMID: 22972300

5. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The Science of Fake News. Science. 2018;. https://doi.org/10.1126/science.aao2998 PMID: 29590025

6. Bode L, Vraga EK. Studying Politics Across Media. Political Communication. 2018; 35(1):1–7. https://doi.org/10.1080/10584609.2017.1334730

7. Hargittai E. Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. The ANNALS of the American Academy of Political and Social Science. 2015; 659(1):63–76. https://doi.org/10.1177/0002716215570866

8. Zannettou S, Caulfield T, De Cristofaro E, Kourtelris N, Leontiadis I, Sirivianos M, et al. The Web Centipede: Understanding How Web Communities Influence Each Other through the Lens of Mainstream and Alternative News Sources. In: Proceedings of the 2017 Internet Measurement Conference. IMC'17. New York, NY, USA: Association for Computing Machinery; 2017. p. 405–417.

9. Potterat J, Woodhouse DE, Muth SQ, Rothenburg RB, Darrow WW, Klovdahl AS, et al. Network Dynamism: History and Lessons of the Colorado Springs Study. In: Network Epidemiology: A Handbook for Survey Design and Data Collection. New York: Oxford University Press; 2004. p. 87–114.

10. De Domenico M, Granell C, Porter MA, Arenas A. The Physics of Spreading Processes in Multilayer Networks. Nature Physics. 2016; 12(10):901–906. https://doi.org/10.1038/nphys3865

11. Gómez S, Díaz-Guilera A, Gómez-Gardeñes J, Pérez-Vicente CJ, Moreno Y, Arenas A. Diffusion Dynamics on Multiplex Networks. Physical Review Letters. 2013; 110(2):028701. https://doi.org/10.1103/PhysRevLett.110.028701 PMID: 23383947

12. Radicchi F, Arenas A. Abrupt Transition in the Structural Formation of Interconnected Networks. Nature Physics. 2013; 9(11):717–720. https://doi.org/10.1038/nphys2761

13. Boccaletti S, Bianconi G, Criado R, del Genio CI, Gómez-Gardeñes J, Romance M, et al. The Structure and Dynamics of Multilayer Networks. Physics Reports. 2014; 544(1):1–122. https://doi.org/10.1016/j.physrep.2014.07.001 PMID: 32834429

14. Kauffman K, Werner CS, Titcomb G, Pender M, Rabezara JY, Herrera JP, et al. Comparing Transmission Potential Networks Based on Social Network Surveys, Close Contacts and Environmental Overlap in Rural Madagascar. Journal of The Royal Society Interface. 2022; 19(186):20210690. https://doi.org/10.1098/rsif.2021.0690 PMID: 35016555

15. Klovdahl AS, Potterat JJ, Woodhouse DE, Muth JB, Muth SQ, Darrow WW. Social Networks and Infectious Disease: The Colorado Springs Study. Social Science & Medicine. 1994; 38(1):79–88. https://doi.org/10.1016/0277-9536(94)90302-6 PMID: 8146718

16. Jalili M, Orouskhani Y, Asgari M, Alipourfard N, Perc M. Link Prediction in Multiplex Online Social Networks. Royal Society Open Science. 2017; 4(2):160863. https://doi.org/10.1098/rsos.160863 PMID: 28386441

17. Morris M, Rothenberg R. HIV Transmission Network Metastudy Project: An Archive of Data from Eight Network Studies, 1988–2001; 2011.

18. Jalili M, Orouskhani Y, Asgari M, Alipourfard N, Matjaž Perc. JOAPP from Link Prediction in Muliplex Online Social Networks; 2017.

19. Watts DJ, Strogatz SH. Collective Dynamics of 'Small-World' Networks. Nature. 1998; 393(6684):440–442. https://doi.org/10.1038/30918 PMID: 9623998

20. Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer Networks. Journal of Complex Networks. 2014; 2(3):203–271. https://doi.org/10.1093/comnet/cnu016

21. Kapferer B. Norms and the Manipulation of Relationships in a Work Context. In: Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns. Manchester University Press; 1969. p. 181–244.

22. May RM, Anderson RM. Transmission Dynamics of HIV Infection. Nature. 1987; 326(6109):137–142. https://doi.org/10.1038/326137a0 PMID: 3821890

23. Watts DJ, Dodds PS. Influentials, Networks, and Public Opinion Formation. Journal of Consumer Research. 2007; 34(4):441–458. https://doi.org/10.1086/518527

24. Valente TW. Network Models of the Diffusion of Innovations. Hampton Press; 1995.

25. Guilbeault D, Centola D. Topological Measures for Identifying and Predicting the Spread of Complex Contagions. Nature Communications. 2021; 12(1):4430. https://doi.org/10.1038/s41467-021-24704-6 PMID: 34285206

26. Granell C, Gómez S, Arenas A. Dynamical Interplay between Awareness and Epidemic Spreading in Multiplex Networks. Physical Review Letters. 2013; 111(12):128701. https://doi.org/10.1103/PhysRevLett.111.128701 PMID: 24093306