2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

SeCoNet: A Heterosexual Contact Network Growth Model for Human Papillomavirus Disease Simulation

1st Weiyi Wang
School of Computer Science, Faculty of Engineering
University of Sydney, Sydney, Australia
Email: Weiyi.Wang@sydney.edu.au

2nd Mahendra Piraveenan
School of Computer Science, Faculty of Engineering
University of Sydney, Sydney, Australia
Email: Mahendrarajah.Piraveenan@sydney.edu.au

***Abstract*—Human Papillomavirus infection is the most common sexually transmitted infection, and can cause serious complications such as cervical cancer. Due to the scarcity of empirical data about sexual relationships in varying demographics, computationally modelling the underlying sexual contact networks is important to understand Human Papillomavirus infection dynamics. Here we present SeCoNet, a heterosexual contact network growth model for Human Papillomavirus disease simulation. The growth model consists of three mechanisms that closely imitate real-world relationship forming and discontinuation processes in sexual contact networks. We also undertake disease dynamics analysis of Human Papillomavirusus using this model.**

I. Introduction

Human Papillomavirus (HPV) is the most common sexually transmitted viral infection [1]. The incidence of Cervical Intraepithelial Neoplasia (CIN) and cervical cancer is attributed to persistent high-risk HPV infection. Several modelling approaches have been utilized to simulate HPV disease dynamics, predict the final HPV epidemic magnitude, and evaluate the efficacy of prevention strategies. In most of these, the population is assumed to be homogeneously mixed, and the heterogeneity of relationship networks is not captured. However, when it comes to sexually transmitted infections in general and HPV in particular, infectivity is not homogeneous, and sexual contact provides the prerequisite route for infection dissemination. Complex network models can represent these sexual contacts whereby a node in the network denotes an individual, and an edge between two nodes denotes the relationship between them. Generally, it is assumed that a node's transmissibility depends on, and is proportional to, its degree i.e., the number of edges it has. Empirical studies suggest that the underlying topology of a contact network can have a profound impact on the epidemic transmission dynamics of HPV, and the threshold for epidemic outbreaks in complex network models can be quite different from that in classical well-mixed or homogeneously mixed models [2]. Classical network models such as lattices, random networks, small-world networks and scale-free networks have been used to represent sexual contact networks [2].

Scale-free topology is ubiquitous in social networks. Empirical studies [3] suggest that a person's sexual attractiveness is positively correlated with people's knowledge of the given person's number of sexual partners, essentially suggesting a rich-gets-richer mechanism [4] in terms of the number of sexual partners. Furthermore, it has been shown that in sexual contact networks, the number of sexual partners for each person is distributed according to a power-law [5], implying that the underlying contact networks are scale-free [6]. Therefore, in this work we propose a sexual contact network growth model, particularly one that is relevant to HPV transmission, that produces scale-free networks by mimicking real-world relationship building processes. We analyse the topological properties of the resultant contact networks, and simulate HPV disease dynamics on the resultant contact networks to verify their utility in analysing HPV transmission. We keep our focus on heterosexual relationships in this work. The sexual contact network growth model proposed here has a number of parameters which can be calibrated to match differing demographics, though the current analysis is based on parameters that match Australian demographics.

It should be noted here that, though there are several types of sexual contacts, we focus here on heterosexual contacts (via vaginal intercourse) only. The majority of cervical cancers result when the human papillomavirus (HPV) is transmitted from a man to a woman during vaginal intercourse [7]. Though women are primarily affected by cervical cancer and other complications resulting from HPV infection, it has been shown that the risk of transmission though women-to-women contacts can be neglected [8]. Therefore, while acknowledging the broad spectrum of sexual orientations and behaviours present in human society, we nevertheless assume that heterosexual interactions via vaginal intercourse are primarily responsible for HPV spread and its more serious consequences [7], [9]. Therefore, the growth model focusses on heterosexual contacts and is presented as a bipartite network growth mechanism, which is a reasonable simplification in the context of HPV. It should not be assumed however that the growth model captures sexual interactions well in the contexts of other diseases.

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
https://doi.org/10.1145/3625007.3627478

II. Methodology

Here we describe the proposed heterosexual contact network growth model and the HPV transmission model used in this work. Both the sexual contact network growth model and the HPV transmission model are implemented using Netlogo [10].

A. The SeCoNet Sexual Contact Network Growth Model

We propose a growth model which will essentially create scale-free networks, with constraints which are specific to a sexual contact network. We focus on heterosexual contacts only in this study, thus the grown networks will be bipartite networks. The justification for creating bipartite scale-free networks to model heterosexual contact networks is as follows. The number of sexual contacts is not distributed homogeneously among people. People have personal attributes, or perceived attributes, such as age, income, appearance, education, social status etc. which make some people more attractive as sexual partners, resulting in a heterogenous distribution in terms of number of sexual contacts. The involvement of career choices such as being sex workers and perspectives on commercial sex also impact the Lifetime Sexual Partners (LSPs) a person will accumulate. Furthermore, flirting skills typically improve with practise, potentially increasing the success probability of pick-ups in return, which corresponds to a rich-gets-richer mechanism [4] which typically results in the creation of scale-free networks. It has also been demonstrated that real world sexual contact networks exhibit scale-free topology [5]. For these reasons, we propose a sexual contact network growth model which will result in the creation of scale-free networks with characteristics particular to heterosexual contact networks. We name this the SeCoNet (**Se**xual **Co**ntact **Net**work) Growth Model, which is described below. The model at present uses attributes typical to the Australian population, but can easily be calibrated to be applicable to other populations as well. For example, parameters proposed and used in Tracy et al. [11] could be adapted to simulate HPV dynamics in Mali using our growth model.

1) Initialisation: The SeCoNet growth model generates a scale-free heterosexual network consisting of N individuals and their relationships, in order to capture the individual heterogeneity in sexual contact dynamics, where the individuals are denoted by nodes, and the relationships are represented by edges in the network. Each individual (node) i is assigned a set of characteristics and attributes at initialisation: (1) age g_i; (2) gender b_i (1 for females and -1 for males); (3) the estimated average duration of relationships δ_i. The individuals are assumed to be aged between 15 to 49 years old. The age distribution is based on the age distribution of Australia as described in the 2021 Australian census [12], which is summarised in Table I.

TABLE I
Age distribution used in SeCoNet Growth Model

Age Group	%	Age Group	%
15 to 19	12.3	35 to 39	15.6
20 to 24	13.4	40 to 44	14.0
25 to 29	15.1	45 to 49	13.8
30 to 34	15.8		

The model has a range of other variables and parameters, which are summarised in Table II. When the model is initialized, it is assumed that individuals under 20 are virgins, and have had no sexual partners. It is assumed that women prefer older men and vice-versa, with an average age difference $\langle \eta \rangle = 3.5$ [16]. The sexual contact network growth model is implemented with three mechanisms: (1) new node introduction; (2) link removals; (3) secondary link formation. The first mechanism represents new people (nodes) entering (becoming part of) the contact network (as opposed to the relevant populations, which they are assumed to be already members of). Naturally, they need to create links with existing nodes in order to become part of the network, and the parameter m represents how many links they make, on average, when they first become part of the sexual contact network. In this respect this mechanism is essentially similar to classical scale-free network growth models such as the Barabasi-Albert model [6] or the Bianconi-Barabasi model [17]. The second mechanism represents discontinuation of relationships between people who are already in the sexual contact network, with the possibility that a person may 'leave' the network temporarily if they have no links (no active relationships) at a particular time. The third mechanism represents the process whereby new links (relationships) are created between existing nodes (people) in the contact network. Herein lies the significance of this model: it does not merely creates a scale-free network, which several existing growth models can do, but it creates a hererosexual contact network by mimicking the real-world process of relationship formation and maintenance, which is unique to sexual contact networks. The growth model also consists of two phases: in Phase 1, all three mechanisms work simultaneously (that is, people are continuing to join the contact network), while in Phase 2, only the last two mechanisms work (whereby it is assumed that all people in the considered population have joined the network, and new relationships are formed only between people (nodes) which have had at least one relationship previously). The following subsections describe the three mechanisms in more detail:

2) Mechanism I - New Node Introduction: Similar to the Barabasi-Albert model and the Bianconi-Barabasi model, in the beginning $t_0 = 0$, there are m_0 bipartite edges connecting m_0 pairs of nodes in the network, representing m_0 monogamous heterosexual relationships. At each time step t, a new node i is introduced to the network and linked to m preexisting nodes, until there is no new node available. The selection of m nodes to be connected to node i is made preferentially, and depends on both degree and fitness of the existing nodes.

The fitness of node j is defined as:

$$\phi_j = \frac{max\{\langle \eta \rangle, abs(g_i - g_j)\} * abs(b_i - b_j)}{max\{1, abs(lsp_i - lsp_j)\}} \tag{1}$$

where $lsp_i = T/\delta_i$ denotes the estimated Lifetime Sexual Partners (LSP) for node i. The pairing depends on node i and j's age, gender and LSP. The fitness of a node which is not of opposite gender is therefore zero. The probability of node i choosing node j is defined as:

$$p_j = \frac{(k_j + \epsilon)\phi_j}{\sum_{h \in N}(k_h + \epsilon)\phi_h} \tag{2}$$

TABLE II
LIST OF VARIABLES AND PARAMETERS FOR THE SECONET GROWTH MODEL

Variable / Parameter	Symbol	Value	Source
Timestep	t		
Number of contacts	M		
Age	g_i		Australia Bureau of Statistics [12]
Gender	b_i	1: female, -1: male, uniform distribution	
Average Relationship duration for the node (in days)	δ_i	Poisson distribution with a mean of $\langle\delta\rangle = 500$	
Lifetime Sexual Partners	lsp_i	T/δ_i (rounded)	
Cumulative sexual partners	sp_i	$[0, lsp_i)$	
HPV-16 recovery period (in days)	α_i	Exponential distribution with a mean of 390 days	Olsen and Jepson [13]
HPV-16 transmissibility per coital act	β	0.3	Olsen and Jepson [13]
Scale-free exponent	γ		[4]
Scale-free fitness	R^2		
Node Degree	k_i		
Age difference between nodes	$\eta_{i,j}$	$g_i - g_j$	
Expected Relationship durartion	$\Delta_{i,j}$		
Relationship age	$z_{i,j}$		
Total timesteps	T	9000	
Population size	N	3000	
Initial number of links	m_0	5	
Frequency of intercourse	f	$1/2$ during the first two weeks $1/7$ after the first two weeks	Althaus et al. [14]
Scale-free calibration parameter	ϵ	0.5	This can be calibrated.
Links added per joining node	m		[4]
Link removal rate	θ	$1/\langle\Delta\rangle$	
Removed links per node per time step	m_r	$\frac{\theta}{N}$	
Secondary links added per node per time step	m_i		(5), (6)
Average degree	$\langle k\rangle$		
Assortativity	r		[15]
Average age difference	$\langle\eta\rangle$	3.5	Conroy-Beam and Buss [16]

where k_j is the current degree of node j, and the existence of ϵ ensures that nodes which have currently no links (due to all previous relationships being discontinued) are still able to be preferentially selected based on their fitness. Therefore ϵ needs to be a positive real number, and it can be assigned a value to calibrate the growth model. Links created through this mechanism are called primary links hereafrer, to distinguish them from links that are created from mechanism III.

3) Mechanism II - Link Removals: Several growth models only consider forming connections; however, in this context, not all relationships can last permanently. Thus, besides relationship formation, we have added a new mechanism for relationship discontinuation.

Therefore when each link is formed, it is assigned with an expected relationship duration, Δ_{ij}. The expected duration of the relationship (edge) depends on the average relationship duration assigned to both partners (nodes), such that:

$$\Delta_{ij} \sim Exp(min\{\delta_i, \delta_j\}) \tag{3}$$

When the age of the link reaches its duration, the link will be removed, demonstrating the termination of the relationship. Therefore, at each time step t, the rate of edges to be removed θ is $1/\langle\Delta\rangle$, and the number of removed edges per node per timestep is written as $m_r = \theta/N$.

4) Mechanism III - Secondary Link Formation: In the context of sexual contact network formation, two individuals who are both already members of the contact network may in time form a relationship with each other, and this does not have to be limited to the time when either of them joins the network. Thus, we propose a third mechanism to create secondary links between existing nodes. We postulate that the creation of such secondary links will be at a rate which matches the deletion of primary links. Therefore, after mechanism I is completed, the number of links will be stable

in the network. This mechanism uses fitness-based preferential attachment similar to Mechanism I to create secondary links between nodes which are already part of the contact network. It is assumed, without loss of generality, that secondary links are initiated by female nodes. The probability of a female node initiating a secondary link (this should not be taken to mean 'second' relationship - the female node may have any number of primary links (relationships) already) is defined as:

$$q_i = \frac{max\{0, (lsp_i - sp_i)\}}{\sum_{h \in N} max\{0, (lsp_h - sp_h)\}} \tag{4}$$

which means that a female node (person) who has the greatest difference between their expected lifetime sexual partners (defined as the inverse of their average relationship period, δ_i), and the cumulative sexual partners they have already had, is the most likely to initiate secondary relationships. The probability of a male node being selected for this relationship follows a fitness-based preferential attachment mechanism, similar to that in mechanism I.

As mentioned, the growth model has two phases. In Phase 1, all mechanisms operate simultaneously. Therefore, at each time step t, the number of secondary links per node m_i is:

$$m_i = \frac{(mt + m_0)\theta}{N} \tag{5}$$

During phase 2, where only Mechanisms II and III operate, the number of secondary links per node per timestep is:

$$m_i = \frac{M\theta}{N} \tag{6}$$

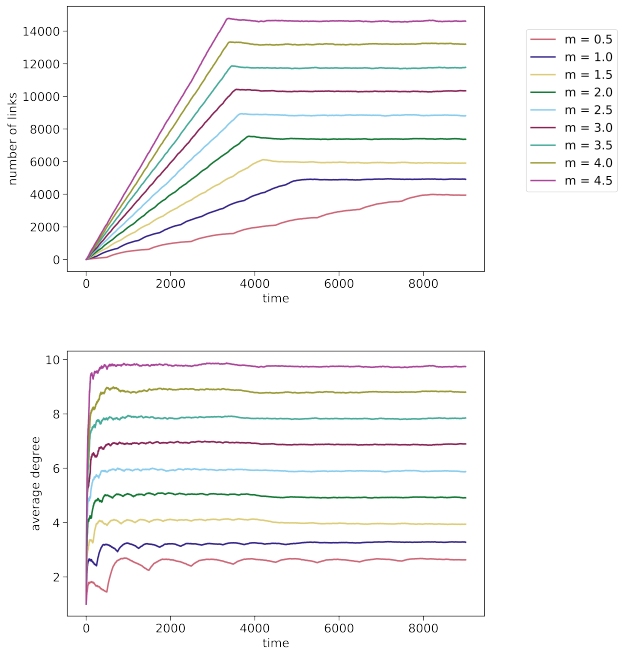

Fig. 1. Number of links and average degree against time in SeCoNet growth model for different m values. Here $N = 3000$, $m_0 = 5$ and $\langle\delta\rangle = 500$.

B. The HPV Transmission Model

The HPV Disease dynamics transmission model is implemented as an SIR epidemic model. The population is compartmentalized into three compartments: (1) Susceptible; (2) Infectious; (3) Recovered, in terms of health status. There is no vaccination implemented in the model, and all recovered individuals are assumed to have lifetime immunity after recovery. Each individual of the population and their sexual contacts are directly represented in the network as nodes and links. The underlying topology represents the heterogeneity in sexual contacts and the transmission dynamics of the infection. In addition to the parameters defined in the sexual contact network growth model, we define epidemic parameters, including population level parameters such as frequency of sexual acts f, transmissibility of HPV-16 per sexual act β, and individual recovery period α_i. These are also included in Table II. As the sexual contact network grows, infection is seeded once for every 100 timesteps: that is, one person out of every hundred people joining the contact network is assumed to be infectious.

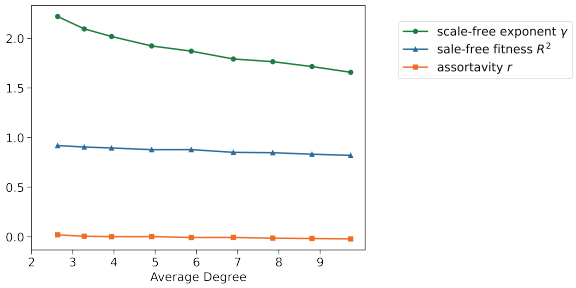

Fig. 2. Topological attributes (scale-free exponent, scale-free fitness, and assortativity) of networks grown by SeCoNetgrowth model. $N = 3000$, $m_0 = 5$ and $\langle\delta\rangle = 500$.

III. Results and Discussion

The simulations described below run for 9000 days. The purpose of these simulations is to demonstrate the suitability of the SeCoNet growth model in simulating HPV disease dynamics. The population N is set as 3000, the initial network consists of $m_0 = 5$ pairs of heterosexual nodes, and the average relationship duration for a node is 500 days (see Table II). In the figures presented below, each data point represents the average of 10 simulation runs unless otherwise stated.

A. Topological Analysis of grown networks

Fig 1 shows the growth of the network (in terms of relationships) for different m values. It can be seen that the network growth speed and the final size of the network is positively correlated with the value of m. Also phase 1 and phase 2 of the SeCoNet growth model are clearly visible, whereby in phase 1 the number of links increase rapidly, while in phase 2 the number of links stabilise. It is clear that in phase 2, the topological parameters change little, even though the actual relationships continue to change. It is also clear that the average-degree of the network at steady state (in phase 2) is closely correlated to the parameter m, as can be expected, even though it is important to note that due to the interplay between the three mechanisms of the SeCoNet model, the parameter m does not directly represent average degree of the resultant network. Further, the higher values of m accelerates the introduction of new nodes, which shortens Phase 1.

Fig. 2 shows the scale-free exponent, scale-free fitness, and degree assortativity [15] of a set of grown networks against average degree. The high scale-free fitness values (ranging from

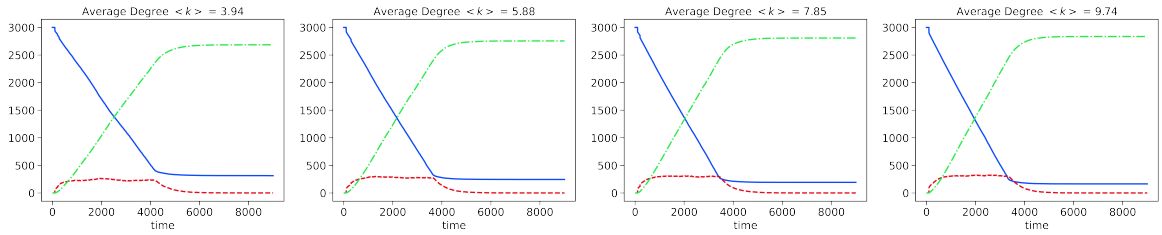

Fig. 3. HPV-16 disease dynamics on networks grown by SeCoNet growth model. Here, transmission probability per sexual act $\beta = 0.3$, mean infection duration $\langle\alpha\rangle = 390$ days. Infection is seeded by introducing one infectious node into the network every 100 timesteps during phase 1.

0.95 to 0.9) demonstrate that the generated networks exhibit scale-free topology. The scale-free exponent γ varies between 2.5 and 1.5, which closely matches the scale-free exponents empirically observed in contact networks [18]. Furthermore, as average degree increases, the scale-free exponent γ decreases. It can be seen that the degree assortativity is close to zero and does not vary with average degree (the nodes do not show preferential mixing in terms of degrees).

B. Simulating HPV Disease dynamics

To confirm that the networks evolved from the SeCoNet growth model are suitable to analyse HPV disease dynamics, we conducted a set of basic infection transmission simulations on these contact networks, as shown in Fig. 3, for a range of networks with differing average degrees. In all-cases, the classic dynamics of SIR models is observed, that is, the susceptible proportion reduces rapidly, the infectious proportion peaks then reduces, and then the recovered proportion increases rapidly, even though the population is not-well mixed and contacts are represented and governed by the underlying contact network topology. It can be observed that the peak amplitude of the infectious proportion increases when average degree increases (the higher connectivity seems to encourage a more vigorous infection transmission). It can also be seen that the network saturates more quickly with recovered people when connectivity increases, due to the higher infection peaks. In all cases though, it could be noted that a very small proportion of the populace remains infected at the end of the simulation, due to the long average recovery period of HPV, which results in some people taking years to recover. In short, all grown networks demonstrate expected disease dynamics.

IV. Conclusion

In this work, we introduce the SeCoNet (Sexual Contact Network) growth model, which produces scale-free contact network models using three mechanisms which reflect the real world relationship forming processes of people. The growth model is focussed on heterosexual relationships and is calibrated to Australian demographics at present, though it can easily be adapted to other demographics. It is intended for simulation of HPV infection. The growth happens in two phases, in the first of which all three growth mechanisms are at work, and the number of relationships rapidly increases, while in the second phase only two mechanisms are at work, and the number of relationships is fairly stable, though the topology continues to change and evolve. Our analysis shows that the model creates highly scale-free networks, and growth parameters can be adjusted to produce networks with a desired scale-free exponent. We demonstrate that a compartmental HPV Infection transmission model can be effectively simulated on networks produced by the SeCoNet growth model.

References

[1] W. H. Organization. Cervical cancer. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer

[2] M. Keeling, "The implications of network structure for epidemic dynamics," vol. 67, no. 1, pp. 1–8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040580904001121

[3] L. A. Dugatkin, *The Imitation Factor: Evolution Beyond the Gene*. Simon and Schuster.

[4] M. Bell, S. Perera, M. Piraveenan, M. Bliemer, T. Latty, and C. Reid, "Network growth models: A behavioural basis for attachment proportional to fitness," vol. 7, p. 42431. [Online]. Available: https://www.nature.com/articles/srep42431

[5] B. F. De Blasio, . Svensson, and F. Liljeros, "Preferential attachment in sexual networks," vol. 104, no. 26, pp. 10 762–10 767.

[6] A.-L. Barabsi and R. Albert, "Emergence of Scaling in Random Networks," vol. 286, no. 5439, pp. 509–512. [Online]. Available: https://www.science.org/doi/10.1126/science.286.5439.509

[7] J. Herbert and J. Coffin, "Reducing patient risk for human papillomavirus infection and cervical cancer," *Journal of Osteopathic Medicine*, vol. 108, no. 2, pp. 65–70, 2008.

[8] A. Burchell, R. Winer, S. de Sanjosé, and E. Franco, "Chapter 6: Epidemiology and transmission dynamics of genital hpv infection. vaccine. 24suppl 3," S3/52–S3/61. doi: 10.1016/j. vaccine. 2006.05. 031, Tech. Rep., 2006.

[9] S. B. Baldwin, D. R. Wallace, M. R. Papenfuss, M. Abrahamsen, L. C. Vaught, J. R. Kornegay, J. A. Hallum, S. A. Redmond, and A. R. Giuliano, "Human papillomavirus infection in men attending a sexually transmitted disease clinic," *The Journal of infectious diseases*, vol. 187, no. 7, pp. 1064–1070, 2003.

[10] U. Wilensky, "NetLogo," Center for Connected Learning and Computer-Based Modeling, Northwestern University. [Online]. Available: http://ccl.northwestern.edu/netlogo/

[11] L. Tracy, H. D. Gaff, C. Burgess, S. Sow, P. E. Gravitt, and J. K. Tracy, "Estimating the impact of human papillomavirus (hpv) vaccination on hpv prevalence and cervical cancer incidence in mali," *Clinical infectious diseases*, vol. 52, no. 5, pp. 641–645, 2011.

[12] Australian bureau of statistics. [Online]. Available: https://profile.id.com.au/australia/five-year-age-groups

[13] J. Olsen and M. R. Jepsen, "Human papillomavirus transmission and cost-effectiveness of introducing quadrivalent HPV vaccination in Denmark," vol. 26, no. 2, pp. 183–191.

[14] C. L. Althaus, K. M. E. Turner, B. V. Schmid, J. C. M. Heijne, M. Kretzschmar, and N. Low, "Transmission of Chlamydia trachomatis through sexual partnerships: A comparison between three individual-based models and empirical data," vol. 9, no. 66, pp. 136–146.

[15] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Information-Cloning of Scale-Free Networks," in *Advances in Artificial Life*, ser. Lecture Notes in Computer Science. Springer, pp. 925–935.

[16] D. Conroy-Beam and D. M. Buss, "Why is age so important in human mating? Evolved age preferences and their influences on multiple mating behaviors," vol. 13, no. 2, pp. 127–157.

[17] G. Bianconi and A.-L. Barabsi, "Competition and multiscaling in evolving networks," vol. 54, no. 4, p. 436. [Online]. Available: https://iopscience.iop.org/article/10.1209/epl/i2001-00260-6/meta

[18] I. Z. Kiss, D. M. Green, and R. R. Kao, "Infectious disease control using contact tracing in random and scale-free networks," vol. 3, no. 6, pp. 55–62. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsif.2005.0079