

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/158226>

How to cite:

Please refer to published version for the most recent bibliographic citation information.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

What is R ? A graph drawer's perspective*

Max Sondag¹, Cagatay Turkey², Sibylle Mohr³, Louise Matthews³, Kai Xu⁴, and Daniel Archambault¹

¹ Swansea University, Swansea, UK
 {m.f.m.sondag,d.w.archambault}@swansea.ac.uk

² Warwick University, Warwick, UK
 Cagatay.Turkey@warwick.ac.uk

³ University of Glasgow, Glasgow, UK
 {Sibylle.Mohr,Louise.Matthews}@glasgow.ac.uk

⁴ Middlesex University, Middlesex, UK
 K.Xu@mdx.ac.uk

1 Introduction

In 2020, the worldwide pandemic of COVID-19 had a profound impact on society. One of the most important metrics that is being used to investigate the effectiveness of various interventions to control the spread of the pandemic is R , the effective reproduction number. It is talked about extensively in the news, and a wide-ranging array of different interventions are put in place by governments with the aim of getting R below 1 to curb the spread of the disease. In this abstract, we will investigate what R means from a graph drawer's perspective and aim to open up interesting and relevant research avenues.

2 Defining R

To define R , we start with a temporal network [2] G of the contacts between people, for example all contacts within a city or country. Each node u represents a person, and a temporal edge $e_t = (u, v)$ represents that person u was in contact with person v at discrete time t . We overlay the disease we are interested in on G . Some nodes will be index cases, the initial cases where the disease emanates from. These nodes are exposed to the disease from outside the network (i.e. a pangolin or international travel). A person v who has become infectious at time t , has a chance to expose their neighbors over edges with a time greater than t until v is no longer infectious. After an incubation time, these neighbors become infectious in turn, and can propagate the disease further. This propagation through the network creates an infection map: A set of rooted directed trees \mathcal{T} , where each node v represents a person and has a value $e(v)$ which indicates when this node was exposed to the disease. Using this infection map, we can define R .

* This work was supported by UKRI EPSRC grant EP/V033670/1 and EP/V054236/1

24 There are two different R 's that
 25 are of interest to epidemiologists.
 26 R_0 (known as, “R naught” or “ba-
 27 sic reproduction number”), which
 28 indicates the capacity for the dis-
 29 ease to spread when the entire
 30 population is susceptible and no
 31 interventions are in place [1], and
 32 the R commonly used by the me-
 33 dia: R_t (“R” or “effective repro-
 34 duction number”), which takes in-
 35 terventions and immunity into ac-
 36 count [3,4]. The t in R_t is the
 37 time point of interest, often a day or a
 38 week, but is ultimately dependent
 39 on the characteristics of the dis-
 40 ease and available data. In graph-
 41 theoretical terms, we can define
 42 R_t as follows. Let Y_t denote the
 43 nodes with $e(v) = t$. Let X_t be
 44 the set of children of Y_t in \mathcal{T} . The
 45 R_t value for a single point in time
 46 is then: $R_t^* = \frac{|X_t|}{|Y_t|}$. Due to prag-
 47 matic issues such as reporting is-
 48 sues and weekly updates for poli-
 49 cymakers, R_t is often averaged
 50 over a period of time (usually 7 days for COVID)

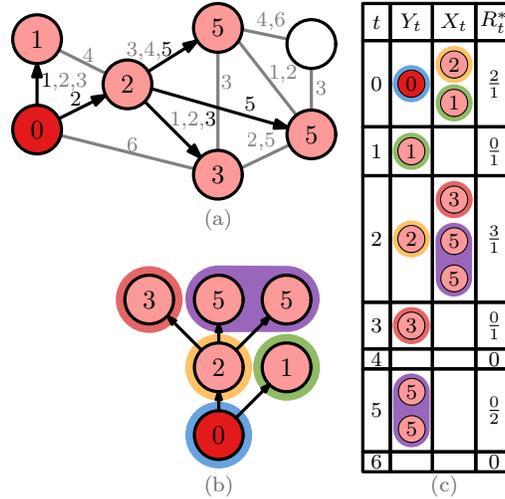


Fig. 1. Example calculation of R_t . (a) Con-
 tact graph with highlighted paths of infec-
 tion. Values within nodes show exposure
 time $e(v)$. (b) The resulting infection map
 is a single infection tree. Nodes with the
 same exposure time have the same color.
 (c) R_t calculation table. R_t is the
 average over the 7 values of R_t^* :
 $(2+0+3+0+0+0+0)/7 \approx 0.71$

51 3 Open problems in real and simulated data

52 We briefly examine various (open) problems
 53 that are encountered when using
 54 this graph-based data. We first consider
 55 problems in the temporal network. In
 56 real data, both edges and nodes are
 57 missing from the graph. The time of
 58 contact could be incorrect, and is at
 59 best an approximation (often at day
 60 accuracy) even for simulated data.
 61 Continuing with problems in determin-
 62 ing the spread of the disease over the
 63 network, the directionality of the edges
 64 is generally unknown. It is also not
 65 typically known for certain whether
 66 one node infected another, or if there
 was an outside influence. The time of
 exposure is often an estimation, as
 testing requires time, and the results
 might even be incorrect. For both
 real and simulated data, these data
 characteristics lead to the graph
 structure of the infection map being
 a large forest of small trees, with a
 few larger trees. While not all of
 these problems and characteristics are
 unique to epidemic graphs, we believe
 that there are nevertheless a number
 of interesting research questions
 present in this setting, and research
 from the Graph Drawing community
 could help in assisting with the
 current pandemic and mitigating
 future pandemics.

References

- 68 1. Anderson, R.M., May, R.M.: Infectious diseases of humans: dynamics and control.
69 Oxford university press (1992)
- 70 2. Holme, P., Saramäki, J.: Temporal networks. *Physics Reports* **519**(3), 97 – 125
71 (2012)
- 72 3. Nishiura, H., Chowell, G.: The effective reproduction number as a prelude to statis-
73 tical estimation of time-dependent epidemic trends. In: *Mathematical and statistical*
74 *estimation approaches in epidemiology*, pp. 103–121 (2009)
- 75 4. Wallinga, J., Lipsitch, M.: How generation intervals shape the relationship between
76 growth rates and reproductive numbers. *Proceedings of the Royal Society B: Bio-*
77 *logical Sciences* **274**(1609), 599–604 (2007)