

Visual methods for exploring multivariate spatio-temporal networks with application to health transport

Confirmation Report

Krisanat Anukarnsakulchularp

Advisors: Dr. Michael Lydeamore and Prof. Dianne Cook

This document contains an interactive element (Fig. 5). For the online version of this report, please visit the following link <https://krisanata-confirmation-report.netlify.app/>

Background

Analysing spatio-temporal network data is a contemporary research problem that has gained increasing interest in the health field, particularly within emergency medical services (EMS) and ambulance transfer systems. Such data capture spatial, temporal, and often multivariate information. The spatial component generally represents geographic locations or spatial geometries, while the temporal component records time-related information through timestamps or time intervals (Rao, Govardhan, and Rao 2012). In addition, the underlying network structure creates connections and multivariate dependencies between locations and transfers. While techniques exist to analyse spatial and temporal components separately, performing analysis, and perhaps more importantly, exploring these components in conjunction with the network structure, remains an open challenge.

Older individuals often require continuous support, including 24-hour care, assistance with daily tasks, and ongoing medical supervision. Thus, many reside in the residential aged care facilities (RACFs), which are specifically designed to provide this comprehensive care (Kearney and Winterbottom 2006). RACFs frequently rely on the ambulance services to facilitate the transfers of an individual to the hospital for both acute emergencies and planned/scheduled medical appointments. This rise in the number of transfers is partly due to population ageing (Harris and Sharma 2018), which puts incredible pressure on emergency medical services, where delay could lead to an increase in health risk (Harmsen et al. 2015). During the COVID-19

pandemic, lockdown measures and movement restrictions further disrupted the delivery of emergency services. The effects of lockdowns and rising transfer demand highlight the need for further analysis to improve the planning and utilisation of ambulance services.

The ambulance transfer data was provided by Ambulance Victoria. The dataset included the following aspects: spatial, temporal, and multivariate information on each transfer. The spatial information covers the location of the aged care facilities and the hospital (destination) in latitude and longitude. The temporal data provides the date of the transfers, and the multivariate information covers hospitals, aged care facilities, and patient-level details.

To gain insight into transfer patterns, data exploration using network representations linking RACFs and hospitals provides a powerful framework. However, most network research focuses primarily on topological properties, often treating the entire network homogeneously and overlooking other important information, such as the association between variables (Cardenas et al. 2021; Fernández-Gracia et al. 2017). While network representation is suited to transfer data, overemphasising network topology can neglect the fundamental principles of data exploration. These limitations arise from the practical challenges of working with spatio-temporal network data, including data cleaning methods, particularly temporal information, the ease of data wrangling and subsetting, and the challenges of visualisation and inference. As a result, simple informative analyses, such as examining variable distributions, temporal trends, or bivariate relationships, are often underutilised, despite the ability to reveal key insights of the data. This underlines the need for an infrastructure that integrates network-based approaches with exploratory data analysis (EDA), enabling a comprehensive exploration of spatio-temporal transfer networks.

Studying how infectious diseases spread throughout the network (transfer between RACFs and hospitals) is important because the older population tend to face a higher risk of mortality during the outbreaks (Parohan et al. 2020). These patient transfers between facilities create ways for the disease to be transmitted across the systems, leading to rapid spread. Traditional compartmental infectious disease models assuming homogeneous or static structure do not adequately capture networks that change over time. In reality, ambulance transfers are highly dynamic, where these connections between facilities can change in response to the demand, constraints, and even outbreak conditions. Understanding these transmission dynamics is therefore crucial for devising effective policies to limit spread as well as identify high-risk facilities, critical transfer connections, and exposed periods.

Project 1: Developing Infrastructure for Exploratory Analysis of Multivariate Spatio-temporal Network with Applications to Ambulance Transfers

Part A: Exploratory Data Analysis Infrastructure for Multivariate Spatio-temporal Network

Multivariate spatio-temporal network data has space, time, multiple variables, associations between different locations (which define the network), and multiple types of relationships (hospital to hospital or aged care to hospital, etc.) As these data become more accessible and complex, understanding their structure and dynamics is key to effective decision-making. A major challenge in analysing this data lies in the sheer amount of information it contains most of which is often overlooked due to the difficulty of switching between these dimensions. This infrastructure aims to support the exploration of multivariate spatio-temporal network data. The exploratory data analysis involves several key processes: data storage, cleaning, subsetting, and visualisation. The following section, therefore, reviews existing tools that support these processes and discusses their limitations.

Data Storage and Cleaning

Data cleaning is the first stage of a reliable analysis. Spatio-temporal data usually need to be checked for inconsistency of the temporal records, duplicated records, and spatial inaccuracies. Now, adding the network structure on top of that, such as nodes, edges, and their attributes, requires preserving the relational structure of the network and the spatio-temporal context together throughout the process. Typically, this stage involves tools such as `dplyr` (Wickham et al. 2023) for manipulating the data, `tsibble` (Wang, Cook, and Hyndman 2020) for validating the temporal inconsistency, `sf` (Pebesma 2018) for checking the coordinate inaccuracies, and `igraph/network` (Csárdi et al. 2026; Butts 2008) for keeping the network structure.

The `tidygraph` (Pedersen 2024b) package provides a tidy API for graph and network manipulation, where network data is thought of as two tidy tables, one for node and one for edge data. In tidy data (Wickham 2014), each variable has its own column, each observation has its own row, and each value has its own cell. These tables are then stored together within a `tbl_graph` object, which preserves the underlying network topology while allowing standard `dplyr` verbs to be applied. The interaction between node and edge tables is done through the use of a special function, `activate()`, which allows the user to switch between the two tables and apply `dplyr` operations such as `mutate()`, `group_by()`, and join operations as shown in Listing 1.

There are two main functions for creating `tbl_graph` object, `as_tbl_graph()` and `tbl_graph()`. The first function `as_tbl_graph()` takes in a different class of objects,

such as `data.frame`, `igraph`, and `network`, then turns it into a `tbl_graph` object. While `tbl_graph()` takes in two `data.frame` objects, one for node and one for edge.

From Listing 1, the difference between these two methods is that for the `as_tbl_graph()` function, it only needs the edges dataset, which means that all the multivariate information will only be on edge data and in the node data, it will only have the name (location). For the `tbl_graph()` function, the node variable can be explicitly stated, which can come in handy when there are attributes on the node dataset.

Listing 1 A function for creating a `tbl_graph` object. There are two main ways: first, using the edge list and second, using node and edge tables. The `dplyr` function will work normally, like it is on a `tibble` object, but requires an `activate()` function to switch between the two tables (node and edge).

```
as_tbl_graph(edges)

tbl_graph(nodes, edges)

graph |>
  activate(edges) |>
  mutate(year = lubridate::year(casedate))
```

```
# A tbl_graph: 815 nodes and 102073 edges
#
# A directed acyclic multigraph with 6 components
#
# Edge Data: 102,073 x 9 (active)
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
#> #> #> #> #> #>
```

```

name          longitude latitude type
<chr>        <dbl>    <dbl> <chr>
1 1 ABERDEEN STREET RESERVOIR      145.   -37.7 racf
2 1 ADENEY STREET CAMPERDOWN      143.   -38.2 racf
3 1 AITKEN AVENUE DONALD         143.   -36.4 racf
# i 812 more rows

```

For spatial networks, the `sfnetworks` package (van der Meer et al. 2024) extends `tidygraph` by allowing spatial geometries to be incorporated directly within the `tbl_graph` object. It is useful for dealing with complex geometry where edges are not straight-line connections, such as road or transport networks. The package also allows for the standard spatial operation within the `sf` package to be performed within the network context.

However, the `tidygraph` object does not provide direct support for time. It would be desirable to use a `tsibble` object as an edge table, which would allow temporal checks, including checking time intervals and missing values. However, in the current framework, these checks have to be performed prior to the creation of the `tbl_graph`. It introduces an important limitation, where common operations such as filling missing observations are done outside the network context and are therefore error prone, and require maintenance of multiple objects.

A basic requirement in cleaning temporal network data is the ability to check whether both nodes exist for a given edge at a specific time. In the case of a missing node, how should the edges associated with that node be imputed? A possible solution is to assume that no edges exist during that period, which may be reasonable in some cases but not in others. It highlights a key challenge in cleaning spatio-temporal network data, where temporal consistency and network structure should be considered jointly. Addressing these requires an object to ensure that both temporal attributes and the relational structure of the network remain coherent throughout the cleaning process.

Data Subsetting

Data subsetting is used to extract a subset of spatio-temporal network data based on spatial, temporal, and multivariate variables. It includes grouping data by time periods or regions, as well as filtering based on variable values and network characteristics (e.g., in-degree). Subsetting is particularly useful for in-depth analysis, where it may need data from all RACFs within a certain distance of a particular hospital or transfers occurring during the weekends.

In a network context, filtering operations need to account for topological dependencies between nodes and edges. When nodes are removed based on a condition, all edges incident to those nodes are also deleted (Figure 1). In contrast, when edges are removed, the nodes connected to those edges are preserved, since nodes can exist independently from an edge (Figure 2). The data subsetting is currently well-handled by the `tidygraph` package. The `tbl_graph` object supports these subsetting operations through the use of `dplyr` functions such as `filter()`

and `select()`, illustrated in Listing 2, which are applied separately on nodes and edges while maintaining the condition of the underlying network. Similarly to the data manipulation, users will need to switch between the node and edge tables to subset based on their attributes.

Listing 2 A filtering method in tidygraph, where the `activate()` function allows the user to switch between the node and edge tables.

```
graph |>
  activate(edges) |>
  filter(between(year, 2020, 2021))

# A tbl_graph: 815 nodes and 48606 edges
#
# A directed acyclic multigraph with 59 components
#
# Edge Data: 48,606 x 9 (active)
  from      to casedate    age gender diagnosis      daytype single_id year
  <int> <int> <date>    <dbl> <chr>  <chr>        <chr>       <dbl> <dbl>
1  522     628 2020-10-03   90 Male   LACERATION    weeken~  11574122 2020
2  562     633 2020-12-18   91 Female  PAIN        weekda~  11813591 2020
3  562     633 2021-03-31   92 Female  PAIN        weekda~  12134872 2021
4  562     640 2021-04-22   92 Female SHORT OF BREATH weekda~  12204603 2021
5  516     706 2021-01-13   98 Female NO PROBLEM IDENT~ weekda~  11895939 2021
6  596     640 2020-07-04   90 Male   ALTERED CONSCIOU~ weeken~  11319542 2020
7  279     769 2020-05-16   19 Male   POST ICTAL      weeken~  11183777 2020
8  231     645 2020-08-06   46 Male   OTHER - SPECIFY weekda~  11422186 2020
9  231     645 2020-08-07   46 Male   OTHER - SPECIFY weekda~  11415518 2020
10 265     657 2021-05-05   45 Male   NO PROBLEM IDENT~ weekda~  12244819 2021
# i 48,596 more rows
#
# Node Data: 815 x 4
  name          longitude latitude type
  <chr>        <dbl>    <dbl> <chr>
1 1 ABERDEEN STREET RESERVOIR      145.   -37.7 racf
2 1 ADENEY STREET CAMPERDOWN      143.   -38.2 racf
3 1 AITKEN AVENUE DONALD         143.   -36.4 racf
# i 812 more rows
```

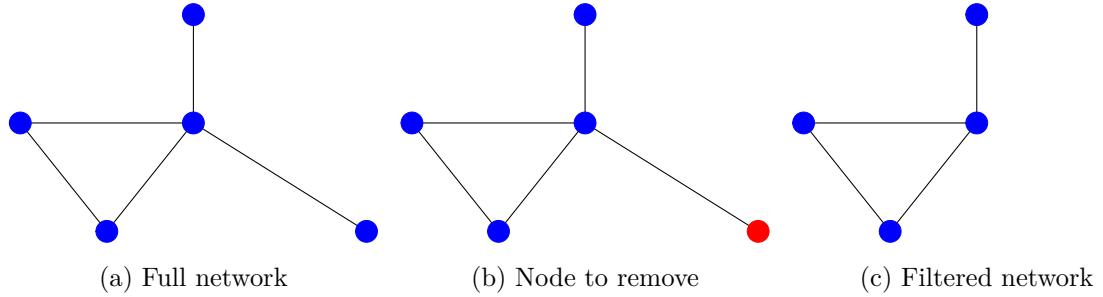


Figure 1: An illustration of node filtering, such as can be done in tidygraph. The selection of a node can be made based on variables such as the facility type (RACF or Hospital). When a node gets filtered out, edges incident to it will also disappear.

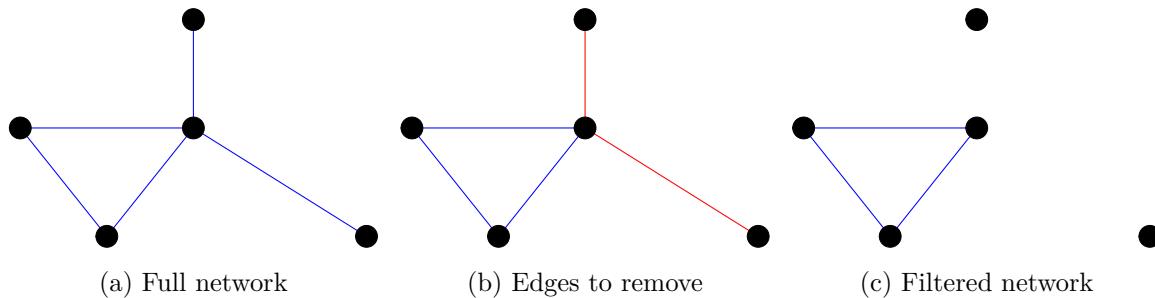


Figure 2: An illustration of edge filtering, such as can be done in tidygraph. The selection of an edge can be made based on variables such as the transfer weekend. When an edge gets filtered out, nodes incident to it will still be kept.

Network Sampling

Another important aspect of subsetting is understanding how sampling methods perform on network data. Observational data are often not evenly distributed across multiple dimensions such as time, space, or variable groups. Some strata may contain more observations than others, and analysing these can directly impact the interpretation, as the larger group of strata may dominate the patterns seen. Sampling provides a way to subset the data while keeping it representative of the population. Stratified sampling, in particular, helps with an imbalance case by dividing the data into subgroups and sampling within each group, ensuring that all groups are represented in the sampled data.

In the network context, sampling methods are generally categorised into the following (Chuong Nguyen 2025):

- **Node-based sampling** selects a subset of nodes from the network and retains edges that are incident to the sampled nodes. This method is efficient and is usually implemented in large-scale studies (Ben-Eliezer et al. 2022). It often fails to capture important global structural properties such as connectivity and clustering.
- **Edge-based sampling** samples a subset of edges directly and includes the nodes incident to those edges. This method is better at preserving structural pattern (Jiao 2024). However, it may introduce bias towards selecting nodes with higher degrees, resulting in biased sampled data.

There are many additional methods for sampling. Hu and Lau (2013) provides a comprehensive survey and taxonomy of graph sampling approaches, which are outside the scope of this project.

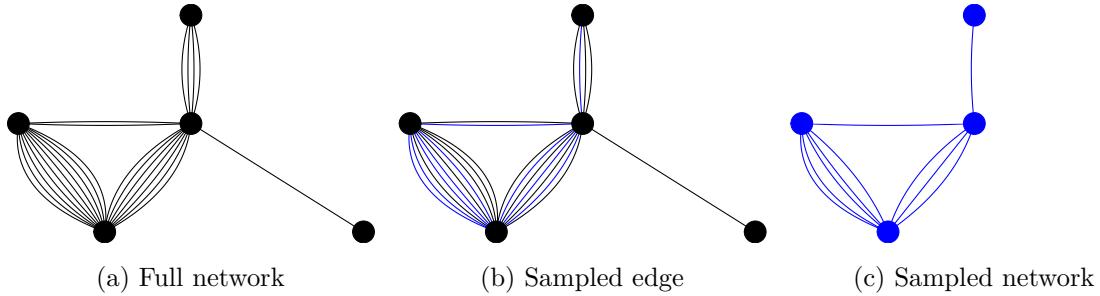


Figure 3: An illustration of edge-based sampling, such as can be done in tidygraph. The edges set will get sampled, then only include the nodes incident to those edges. This method will be biased towards selecting nodes with higher degrees.

The `tidygraph` package provides a method for sampling the data for a `tbl_graph` object through a `sample_n()` function. A further limitation of the `tbl_graph` is that it does not directly support stratified (i.e., `group_by`) sampling. Instead, the `tbl_graph` object needs to

be converted back to `tibble` (Müller and Wickham 2025), performing stratified sampling on the node or edge table, and then filtering the original network based on the sampled nodes or edges (Listing 3). This limitation shows that sampling operations for network objects can still be improved.

Listing 3 A code comparison between sampling and stratified sampling on edges. The tidygraph supports edge sampling using the `sample_n()` function, but with stratified sampling, it does not. This makes it much harder to code stratified sampling.

```
set.seed(1)

# Edges sampling
graph |>
  activate(edges) |>
  sample_n(size = 20) |>
  activate(nodes) |>
  filter(!node_is_isolated())

# Stratified edges sampling
edges_kept <- graph |>
  activate(edges) |>
  as_tibble() |>
  group_by(daytype) |>
  sample_n(size = 10) |>
  pull(single_id)

graph |>
  activate(edges) |>
  filter(single_id %in% edges_kept) |>
  activate(nodes) |>
  filter(!node_is_isolated())
```

```
# A tbl_graph: 33 nodes and 20 edges
#
# A rooted forest with 13 trees
#
# Node Data: 33 x 4 (active)
  name      longitude latitude type
  <chr>     <dbl>    <dbl> <chr>
1 120 MCCRACKEN STREET ESSENDON       145.   -37.7 racf
2 2 CLARKE STREET ABBOTSFORD         145.   -37.8 racf
3 220 MIDDLEBOROUGH ROAD BLACKBURN SOUTH  145.   -37.8 racf
```

```

4 23 FOREST DRIVE FRANKSTON NORTH           145. -38.1 racf
5 242 JELLS ROAD WHEELERS HILL             145. -37.9 racf
6 27 SHIERLAW AVENUE CANTERBURY            145. -37.8 racf
7 27 VICTORIA STREET ELSTERNWICK           145. -37.9 racf
8 359 NARRE WARREN NORTH ROAD NARRE WARREN NORTH 145. -38.0 racf
9 41 LANDSBOROUGH STREET WARRAGUL          146. -38.2 racf
10 45 SILVAN ROAD WATTLE GLEN              145. -37.7 racf
# i 23 more rows
#
# Edge Data: 20 x 9
  from      to casedate    age gender diagnosis      daytype single_id year
<int> <int> <date>    <dbl> <chr> <chr>      <chr>     <dbl> <dbl>
1    15      21 2022-04-04    65 Male   NO PROBLEM IDENTI~ weekda~ 13309866 2022
2     1      24 2020-08-22    97 Female FEBRILE        weeken~ 11456547 2020
3     9      30 2020-12-20    51 Female PSYCHIATRIC EPISO~ weeken~ 11821121 2020
# i 17 more rows

```

The edge-based sampling method will only retain the nodes that are incident to the sampled edges (Figure 3). However, nodes can exist independently without any incident edges, which is what `tidygraph` does for sampling, as discussed in Section . To remove such isolated nodes, they must be explicitly filtered from the node table using the `node_is_isolated()` function. It is also worth mentioning that retaining all nodes, including the isolated ones, after the edge-sampling is possible and is, in fact, the default behaviour in `tidygraph`.

Data visualisation

Data visualisation helps reveal patterns, anomalies and relationships that may not be apparent from numerical summaries alone. Network data is often viewed as connections or flows between nodes/locations, and network-based visualisation allows for easier communication to a broader audience. For a simple network without spatial coordinates, placing nodes and edges in a visualisation requires the use of a graph layout algorithm, such as the Kamada-Kawai layout (Kamada and Kawai 1989). Depending on the chosen algorithm, the positions of nodes and edges can be different even on the same network dataset. With spatial information, visualising these becomes more straightforward, as longitude and latitude can be used to specify the actual location of the nodes, with edges represented as lines connecting these locations.

Visualising multivariate spatio-temporal network data can be challenging. When spatial information is available, the decision to incorporate it into the visualisation is important. On one hand, spatial information allows nodes to be mapped to geographic space, making it easier to infer where nodes are located. On the other hand, omitting spatial information can help reveal non-spatial patterns in the network. Other problems arise when data are highly concentrated in urban areas, making it hard to perceive patterns, or when temporal information

is added. A common approach for handling time is facetting, but comparing them requires many back-and-forths. Another possible solution is to animate the network visualisation.

```
simple_graph |>
  ggraph(x = long, y = lat) +
  geom_sf(data = vic_map, color = "white") +
  geom_edge_link(alpha = 0.1) +
  geom_node_point(aes(color = category)) +
  scale_color_brewer(name = "Facility", palette = "Set1")
```

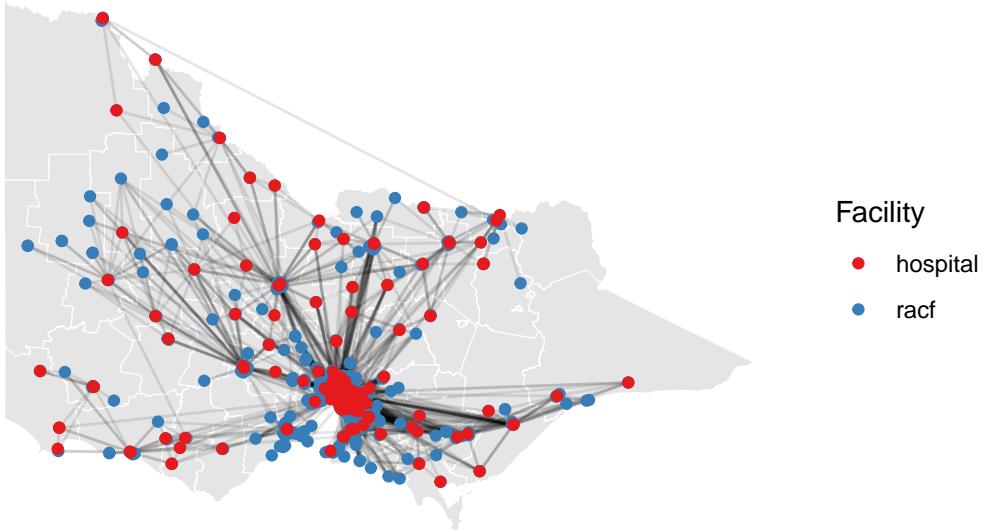


Figure 4: An ambulance transfer network in Victoria between residential aged care facilities and hospitals. The nodes are mapped to the spatial location (latitude and longitude), while edges are the connections between nodes. It can be difficult to observe patterns, since most of the nodes are located within the Melbourne area.

The main limitation of static network visualisation is the restricted amount of information that can be shown in a single figure. The current tool for network visualisation in R is the **ggraph** package (Pedersen 2024a), which extends the **ggplot2** package (Wickham 2016) to support relational data structures such as networks, graphs, and trees. The **ggraph** package is effective at visualising static networks, offering a range of layout algorithms for placing the node locations while keeping the same familiar **ggplot2** syntax. The support for interactive network visualisation with **ggraph** is currently limited. The reason static network visualisation is hard

is that the amount of information that can be mapped to the visualisation is limited within a single figure. As shown in Figure 4, just a simple network representation can already become cluttered quickly. Answering detailed questions such as the number of transfers between a specific RACF and Hospital, or the name of a particular RACF, is difficult using static visualisation alone. Interactive visualisation helps with these limitations by layering additional information onto the visualisation, allowing for further exploration.

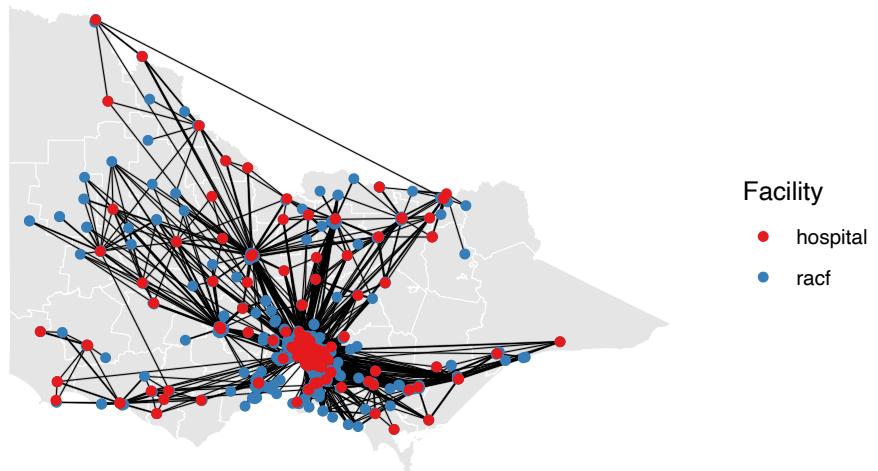


Figure 5: An interactive ambulance transfer network in Victoria between residential aged care facilities and hospitals. The extra information added to the nodes is the location name, and for edges, the number of transfers. This information will be visible when hovering over the node or edge. Using interactive makes it easier to explore the network without adding too much information to the plot.

The `ggiraph` package (Gohel and Skintzos 2025) provides an interactive element to the visualisation through the `ggplot2` extension. Creating interactive network visualisations, therefore, typically requires combining `ggraph` for layout construction with `ggiraph` for interactivity. Getting this to work together requires an insight into how `ggraph` works, which is not a seamless implementation. For example, `ggraph` provides `geom_node_*` and `geom_edge_*` functions for nodes and edges geometry, respectively, but these do not natively support interactivity. To make nodes interactive, `geom_point_interactive()` needs to be used instead of `geom_node_point()`. Making edges interactive is more complex as it requires the use of `geom_segment_interactive()` and mapping the start and end coordinates (`x`, `y`, `xend`, `yend`) for each edge (Listing 4). By that point, if both node and edge need interactivity, then `ggraph` is not needed, and instead the node and edge can be treated as a different dataset and implement `geom_point` and `geom_segment` for drawing the node and edge.

Next steps

- The temporal data are not easy to clean currently with `tidygraph`. As a result, issues such as filling missing timestamps and handling inconsistent time periods are currently handled outside the network context, which may affect the network's validity. Therefore, the new data object needs to integrate temporal and network topology jointly.
- The current data sampling method is limited, with most methods working outside the `tidygraph` object. Although many tools already work with `tibble`, its functionality requires writing a method for `tbl_graph` because `tidygraph` converts `igraph` into `tibble`. Therefore, the new data object should be built on top of `tibble`.
- Data visualisation capabilities are constrained by the `ggraph` package, as layout calculations are handled internally, making it difficult to access and reuse layout information when extending the visualisation method (e.g., interactive visualisation). Therefore, the new data object will be designed to work directly with the `ggplot2` package for better expansion and application of additional geoms for network visualisation.
- The work will result in an R package that provides a data object that better encapsulates the spatio-temporal network structure, along with the necessary functionality. It will be designed to work with a tidyverse suite. This is expected to be written into a publishable paper.

Part B: An Exploratory Analysis of Ambulance Transfer in Victoria

Ambulance and emergency medical services play a critical role in transferring older individuals from residential aged care facilities (RACFs) to hospitals. These transfers form links between RACFs and hospitals across space and time, and are associated with multiple variables such as age, sex, and reason of transfer. Exploratory analysis provides an initial step in examining ambulance transfer data by revealing patterns, anomalies, and relationships. This data can

be represented as a multivariate spatio-temporal network, where facilities are represented as nodes and ambulance transfers as time-varying edges.

During the COVID-19 pandemic, government-imposed lockdowns created significant challenges for ambulance services. Previous studies have shown that the pandemic affected the transfer of older individuals from residential aged care facilities to hospitals (Wyer et al. 2024; Nair et al. 2023; Botan et al. 2023). These restrictions limited ambulance availability and highlighted potential vulnerabilities in the reliance on ambulance transfers for aged care residents. Examining ambulance efficiency and transfer volumes during lockdown periods can inform strategies to improve system resilience and preparedness for future emergencies.

Transfer distance

To quantify the ambulance transfer, instead of just looking at the number of transfers, the distance travelled from the aged care facilities to the hospital is considered. The Euclidean distance will be used to calculate a transfer distance. It is the shortest distance between two points based on the latitude and longitude. This measure should give a reasonable estimate for less computing compared to the actual road distance. Note that the limitation is that it ignores the actual street network distance, which is not a straight line, and therefore, the distance between the RACF and the hospital will generally be longer.

From Figure 6, the median is lower for the emergency transfer compared to scheduled transfers, but not much difference in terms of their distribution. However, there are transfers with zero distance in the data, which might suggest either a data entry error or that the aged care facilities and hospitals are co-located (i.e., located next to each other). In the latter case, this might suggest that the transfer may not be necessary. After filtering the distance data to zero and going through them manually, it can be seen that the aged care facility and hospital are indeed in the same location.

Generally, transfers of the patient should be done quickly, which means sending them to the closest hospital available. It could also explain why some RACFs have a zero transfer distance. Though when the co-located hospital cannot handle emergency cases, the ambulance will need to transfer patients to the nearest hospital with an emergency department. It can be verified using the hospital data, which includes a column indicating whether the hospital has an emergency department. To further analyse this, the distances will be calculated between RACFs and hospitals across all possible combinations to identify which RACFs are located next to hospitals. After the calculation, the RACFs will be categorised into two groups: RACFs close to hospitals and those otherwise.

The expectation of the transfer distance for the co-located RACF would be shorter than that of the other RACFs. However, as shown in Figure 7, the median transfer distance for the co-located RACF is higher. From Figure 8, it can be seen that there are many more co-located RACFs in the regional area than in the Melbourne area, which might explain the bimodality

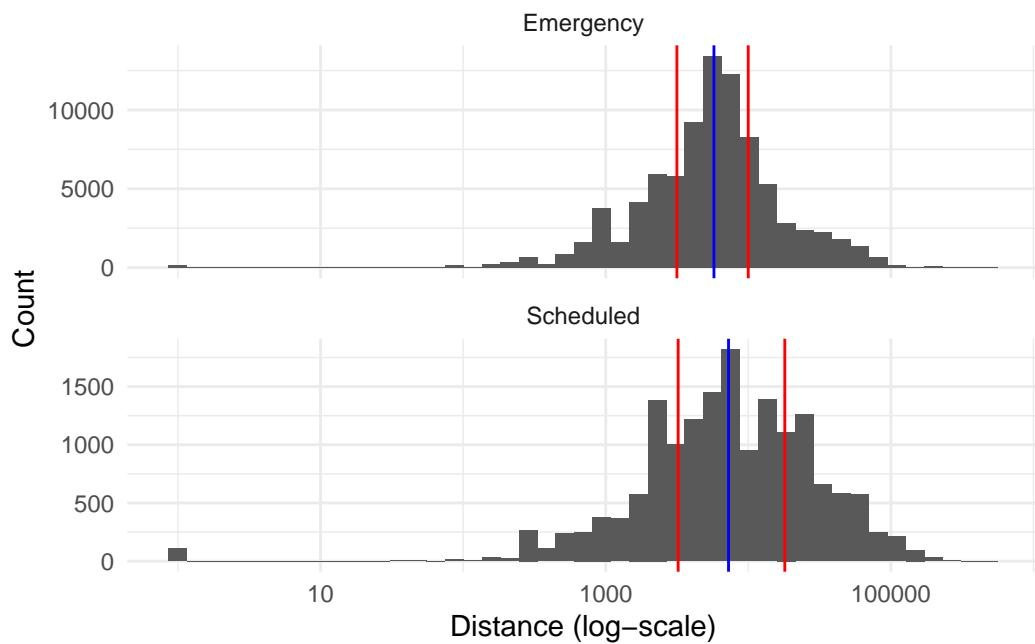


Figure 6: Histogram of ambulance transfer distances for emergency and scheduled cases, with the median indicated by the blue line and the first and third quartiles shown in red. The distributions for emergency and scheduled transfers are similar, with no large difference between the two groups.

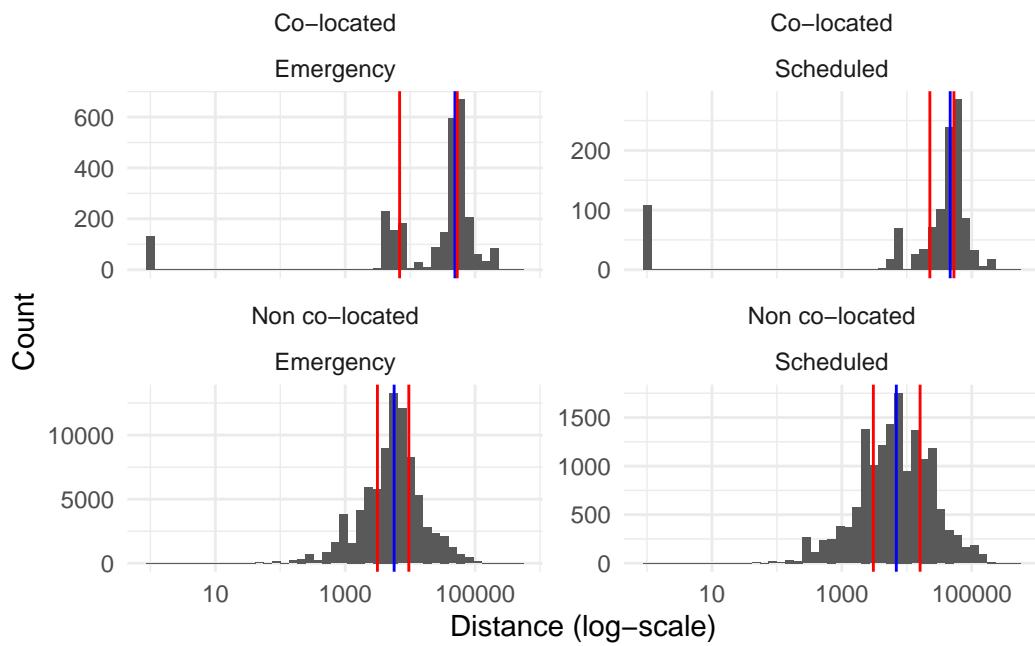


Figure 7: Histogram of ambulance transfer distances for emergency and scheduled cases grouped by co-located RACF. The median is shown in a vertical blue line, while the first and third quartiles are shown in red. The median transfer distance for the co-located RACF is much higher than that of others.

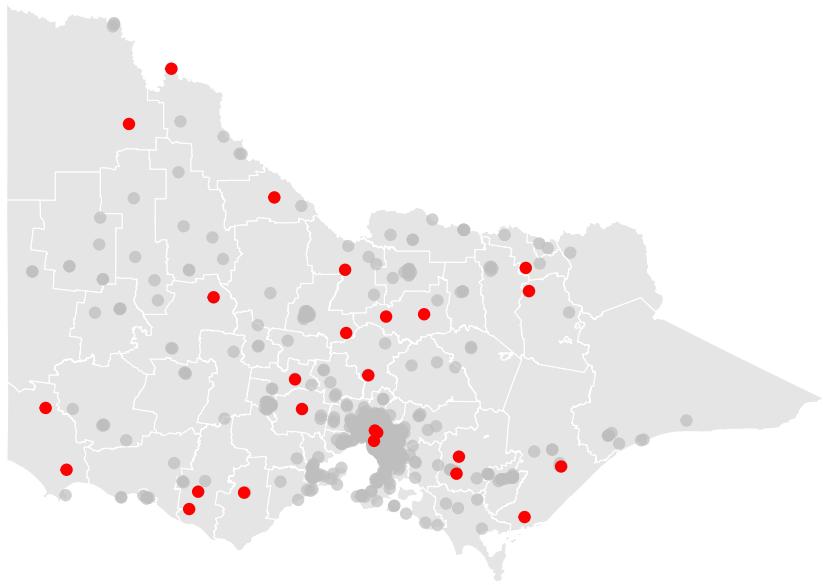


Figure 8: The location of the RACF with the co-located RACF highlighted. The plot shows where all the RACFs are located in Victoria, with the one highlighted as the co-located facility. Most of the co-located facilities are in the regional area, with only a few in the Melbourne area.

observed in Figure 7. This bimodality is likely caused by the longer travel distances for regional RACFs compared to those in Melbourne.

To highlight this regional effect, a sample of RACFs will be visualised on a map. The sampling is used to reduce edge clustering while still representing the observational data, making it easier to observe patterns. Additionally, since there are many facilities inside the Melbourne area, hospitals in Melbourne will be grouped based on their emergency capabilities.

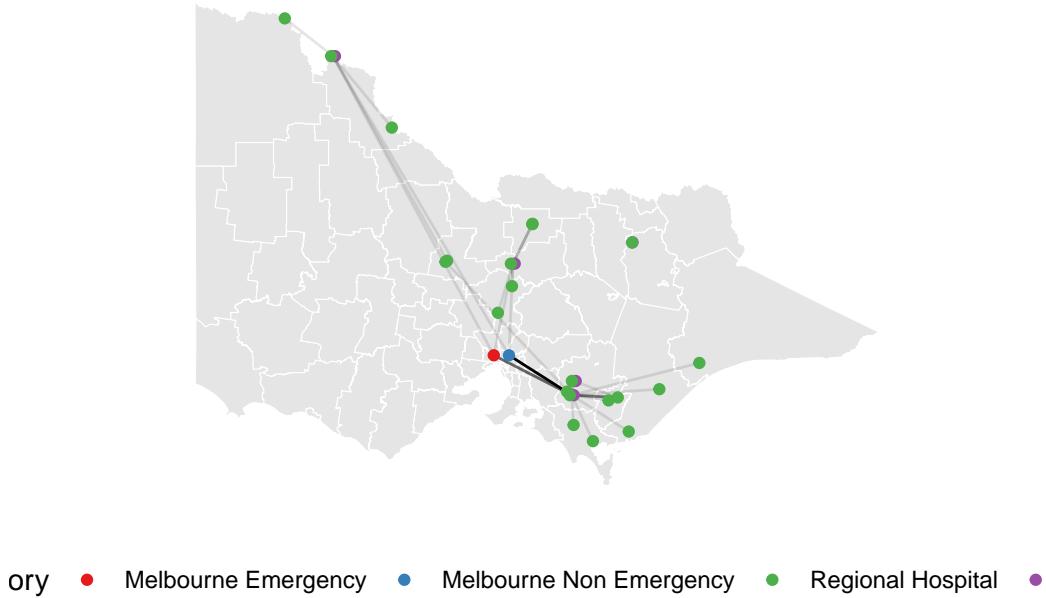


Figure 9: The sampled co-located RACF transfer network, where the Melbourne hospital is grouped. These facilities are placed according to their spatial coordinates (latitude and longitude), with a line representing the transfers and alpha representing the transfer volume. Most patients are sent to Melbourne hospitals.

From Figure 9, it can be observed that the longer median travel distances shown in Figure 7 are caused by the regional RACFs located in the regional area. Using spatial coordinates helps identify this pattern. However, this example also shows why mapping a network directly onto its spatial coordinates may not always be ideal. If the main focus of the plot is to highlight that most patients are transferred to Melbourne hospitals, this is not immediately clear in the spatial visualisation. Longer edges tend to attract more visual attention, and overlapping edges further obscure the transfer patterns.

In contrast, Figure 10 shows the same network but using a graph-based layout instead. This representation makes the pattern of transfers to Melbourne hospital much clearer, and the two clusters are more distinctly separated. This comparison highlights the importance of choosing between spatial and graph-based layouts.

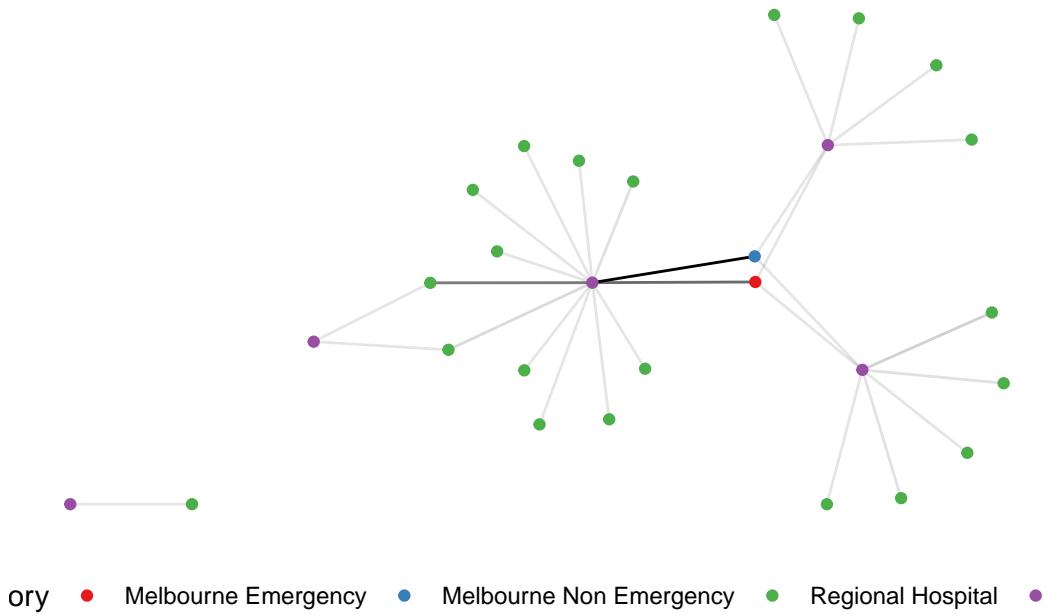


Figure 10: The sample co-located RACF transfer network, where the Melbourne hospital is grouped on the non-spatial layout. The nodes are placed based on the stress layout, which helps spread out the nodes. The transfer of the patient from regional RACFs to Melbourne is clearer, and there are two clusters of transfers.

Extending this to interactive visualisation is not straightforward. In the case of spatial plots, node locations are given in the data, making it easier to add interactivity using functions such as `geom_point_interactive()`. For a graph layout plot, while nodes can be made interactive, the edges are more difficult. It is because the layout positions are calculated internally within `ggraph`, making it challenging to extract and reuse layout information for interactive edges using functions such as `geom_segment_interactive()`.

Following this, transfer reasons, such as dispatch type and diagnosis, will be explored. To make comparison and interpretation easier, the transfer distance will be grouped into four categories: zero-distance (co-located facilities), short-distance (between 0 and 10km), medium-distance (between 10km and 50km), and long-distance (above 50km).

Next steps

- Add road distance to the dataset as an additional distance comparison. Also, including an indicator for aerial transfers can be helpful since aerial transport typically takes less time to transfer the patient to the hospital.
- Conduct a more in-depth analysis of the differences between emergency and scheduled transfers. The initial result (Figure 6) shows that there is not much difference between emergency and scheduled in terms of descriptive statistics and distribution. Further analysis would help to see if there are potential confounding variables.
- Analyse the potential overuse of ambulance callouts. Ideally, the resources should not be overused in non-emergency situations, identifying potential overuse would allow for better resource allocation when needed (e.g., emergency state).
- Examine the temporal patterns, such as peak demand for transfers. It would help inform policymakers about when additional ambulance resources are needed to prevent a system shortage.
- Assess the impact of COVID-19 on transfer patterns. Understanding how COVID-19 affected the system would support future planning in the event of another pandemic or disruption.
- Consider the implications for policymakers, such as the effects of a hospital shutdown. During the disease outbreaks, some hospitals may need to reserve their resources for specific cases, or the hospital is at full capacity. Understanding how this affects the transfer network would be useful for planning.
- The work of Part B will help inform the development of the new data package, as well as applying the new package to the analysis. The analysis of the ambulance transfers will be turned into a publishable application paper.

Table 1: The dispatch reason proportion table for different distance categories. These tables display the top seven common dispatch reasons for each category. The common dispatch reason for zero and medium distances is urgent transfers, while medium to long distances is mental health.

(a) Zero Distance	
Dispatch	Proportion
AMBULANCE-URGENT WITHIN 25 MINS	0.14
AMBULANCE-CRITICAL	0.08
DR REQUESTING ATTENDANCE WITHIN 25 MIN	0.07
MENTAL HEALTH: ACUTE PROBLEMS	0.05
NEPT 000 EVENT TO ERTCOMM	0.05
NEPT BOOKED EVENT TO ERTCOMM	0.05
CLINICIAN-ACTIONED MEDIUM ACUITY EVENT	0.04
(b) Short Distance	
Dispatch	Proportion
AMBULANCE-URGENT WITHIN 25 MINS	0.15
DR REQUESTING ATTENDANCE WITHIN 25 MIN	0.12
CLINICIAN-ACTIONED MEDIUM ACUITY EVENT	0.09
EMERGENCY DEPT TRANSFER - ON DAY	0.08
NO CONSENT-NOT URGENT/NO ASP 1-HOUR	0.05
AMBULANCE-CRITICAL	0.05
UNCONSCIOUS/FAINTING, NOT ALERT	0.04
(c) Medium Distance	
Dispatch	Proportion
MENTAL HEALTH: NON URGENT	0.14
MENTAL HEALTH: ACUTE PROBLEMS	0.11
AMBULANCE EMERGENCY MULTILEG	0.10
AMBULANCE-URGENT WITHIN 25 MINS	0.09
EMERGENCY DEPT TRANSFER - ON DAY	0.07
CLINICIAN-ACTIONED MEDIUM ACUITY EVENT	0.06
NEPT BOOKED EVENT TO ERTCOMM	0.06
(d) Long Distance	
Dispatch	Proportion
CLINICIAN-ACTIONED MEDIUM ACUITY EVENT	0.16
MENTAL HEALTH: ACUTE PROBLEMS	0.13
MENTAL HEALTH: NON URGENT	0.10
AMBULANCE EMERGENCY MULTILEG	0.08
CLINICIAN-ACTIONED LOW ACUITY EVENT	0.05
EMERGENCY DEPT TRANSFER - ON DAY	0.05
AMBULANCE-URGENT WITHIN 25 MINS	0.05

Table 2: The diagnosis reason proportion table for different distance categories. These tables display the top seven common diagnosis for each category. The common diagnosis for zero and medium distances is pain, while for medium to long distances is a psychiatric episode.

(a) Zero Distance		(b) Short Distance	
Diagnosis	Proportion	Diagnosis	Proportion
PAIN	0.09	PAIN	0.13
FRACTURE/S	0.08	OTHER - SPECIFY	0.09
OTHER - SPECIFY	0.08	SEPSIS	0.05
NO PROBLEM IDENTIFIED	0.07	FEBRILE	0.04
PSYCHIATRIC EPISODE	0.07	ALTERED CONSCIOUS STATE	0.04
SEPSIS	0.05	FRACTURE/S	0.03
LACERATION	0.05	INFECTION - OTHER / NOT LISTED	0.03
(c) Medium Distance		(d) Long Distance	
Diagnosis	Proportion	Diagnosis	Proportion
PSYCHIATRIC EPISODE	0.21	PSYCHIATRIC EPISODE	0.16
OTHER - SPECIFY	0.10	OTHER - SPECIFY	0.13
PAIN	0.10	PAIN	0.08
SUICIDAL IDEATION	0.07	SUICIDAL IDEATION	0.07
FRACTURE/S	0.06	FRACTURE/S	0.05
UNKNOWN PROBLEM	0.04	ACUTE CORONARY SYNDROME	0.04
NO PROBLEM IDENTIFIED	0.03	UNKNOWN PROBLEM	0.03

Project 2: Dynamics of Infectious Disease Modelling using a Generalised Ambulance Model

Part A: Generalised Ambulance Model

This project focuses on understanding the impact of any changes in the ambulance transfer network on the spread of infectious diseases, particularly among older populations in residential aged care facilities. Ambulance transfer creates a way in which infections can be transmitted throughout facilities and hospitals (Gruber et al. 2013). Variations in transfer volume, patterns, or constraints can alter the structure of the network and, in turn, influence outbreak dynamics.

There has been research using a spatio-temporal point process model to study ambulance demand rather than modelling the network structure. However, such work has primarily focused on ambulance deployment planning rather than transmission or transfer networks (Zhou et al. 2015). Since the main analysis is to examine the transfer patterns, models such as this do not adequately represent the connections between facilities, but instead focus on predicting demand over the area. Understanding the underlying contact network is therefore what is needed for studying the transfer network connections (Silk et al. 2017). There is a wide range of statistical network models available for studying, and these are potential models to look at.

- **Exponential Random Graph Model (ERGM):** An ERGM is used to study the structure of network data by modelling the probability of observing a given network as a function of network structure and the characteristics of individuals (nodes) within the network (Robins et al. 2007). An advantage of ERGM is their ability to include individual traits (e.g., sex and age), which can be used to explain connection patterns and the likelihood of interacting with similar individuals. A limitation of the model is its lack of flexibility in specifying interaction terms, which is possible but not straightforward. ERGM can be fitted using the `ergm` R package (Pavel N. Krivitsky et al. 2023). For dynamic networks, temporal extensions such as the temporal exponential random graph model can be fitted using the `btergm` package (Leifeld, Cranmer, and Desmarais 2018).
- **Latent Space Model:** Latent space models provide an alternative method to ERGMs for modelling relational data. These models basically work like generalised linear models for edge values, while controlling for the network dependence by placing nodes in k-dimensional space according to their social network distance (Hoff, Raftery, and Handcock 2002; Pavel N. Krivitsky et al. 2009). The advantage of this model is its simpler implementation and fitting compared to ERGMs. However, interpretation of model coefficients can be challenging if the position of nodes in latent space covaries with values of nodal attributes (Cranmer et al. 2017). Latent space model can be fitted using the `latentnet` package (Pavel N. Krivitsky and Handcock 2008).

Part B: Dynamics of Infectious Disease Modelling

Antimicrobial resistance (AMR) poses a significant threat to our health. Several studies show that nursing homes or long-term care facilities are considered an important breeding ground for AMR (Tandan et al. 2022; Strausbaugh et al. 1996; John Jr and Ribner 1991). Therefore, RACFs act as important reservoirs for AMR, creating a concern for the broader healthcare system. Transfers of patients through ambulance between RACFs and hospitals create the pathways for these resistant organisms to spread around the system. Understanding patient movement patterns and characterising the spread is therefore essential for monitoring and controlling the transmission.

Infectious disease dynamics describe how the disease is spread and evolves within the population over time. These dynamics depend on the contact between individuals or groups, movement between locations, and the timing of infection and recovery. Traditional compartmental models, such as susceptible-infectious-recovered (SIR) or susceptible-exposed-infectious-recovered (SEIR), are often used to represent the disease progression within and between groups. Through this modelling, policymakers can conduct scenario analyses on the ambulance transfer network. For example, these models can be used to assess how an outbreak evolves if a particular hospital becomes unavailable due to an outbreak and how this disruption affects disease transmission across the system.

Using a simulated network as input to an infectious disease transmission model allows disease spread to occur both within facilities (RACFs and hospitals) and between facilities through ambulance transfers. Similar approaches have been explored in the literature. However, rather than using the simulated network, it uses a point process model to predict demand and then uses this output as input to an infectious disease model (Amaral, González, and Moraga 2023).

Timeline

Progress Update

Project 1A) Data cleaning and subsetting

- I have explored multiple network analysis software, including `tidygraph`, `igraph` and `network` for comparison (Section). This helps identify the limitations of the current tools and what is needed for the new data object and its functionality.
- I have applied the above software on [Caribou dataset](#), with the main focus on the temporal pattern. It helps see that the limitations of the `tidygraph` are its lack of support for temporal variables. It leads to considering an imputation strategy for temporal network data since there are many time gaps (Section) in the Caribou dataset that might affect the distance analysis.

Project 1A) Data visualisation

- I have experimented with crosstalk to create linked, side-by-side visualisations to combine a network and a simple descriptive plot. However, this will not be used going forward.
- I constructed a supervisor-student relationship network as part of learning about network theory. This dataset is helpful since there are no strict spatial coordinates, thus, experiment with graph layout for placing the nodes.
- I explored `ggiraph` for interactive network visualisation (Section). When drawing a large network, the amount of information shown on the plot can be limited, therefore, looking at alternatives such as interactivity to add more information to the visualisation.

Project 1B) Exploratory data analysis

- I did the preliminary analysis of the ambulance transfer efficiency (Section). This would help inform the development of the new data object and its functionality package.

Project 2A)

- I try a statistical network model, such as `ergm` and `intergraph` for switching between `igraph` and `network` objects. It helps inform what the research gaps are in terms of statistical modelling with network data.

Timeline

Potential papers output

- Paper 1: Methods for exploring multivariate spatio-temporal networks: A Review.
- Paper 2: Understanding the spatio-temporal network dynamics on aged care ambulance transfers in Victoria.
- Paper 3: A new tidy data structure for spatio-temporal multivariate network exploration and modelling.
- Paper 4: Modelling infectious disease dynamics using a generalised ambulance model.

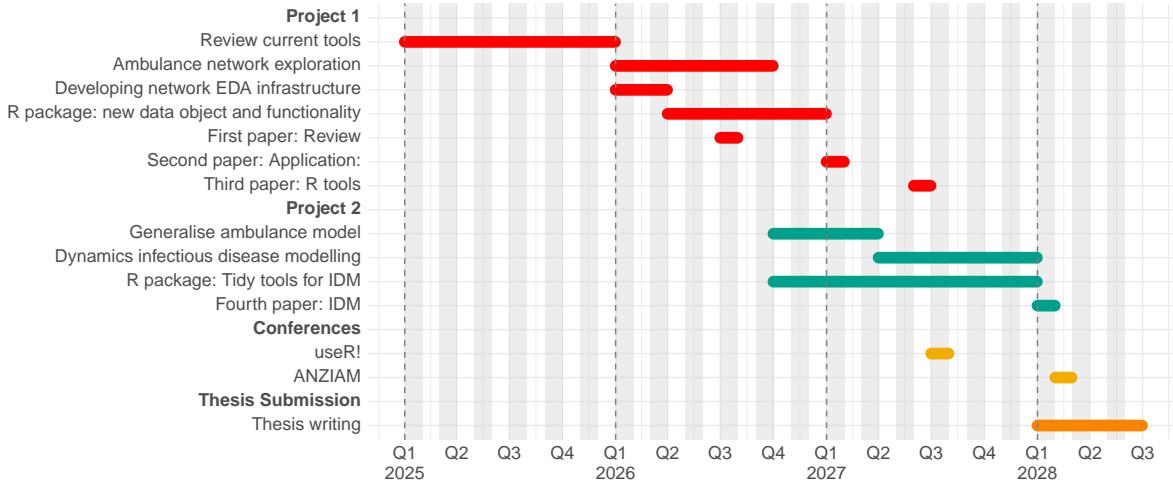


Figure 11: A Gantt chart of my PhD project timeline. The sub-project name is shown alongside the expected timeline. Project 1 is planned to finish by the end of this year, and Project 2 is planned to finish by the end of next year.

Planned conferences

The following is a list of conferences that I will be targeting to attend and showcase the work.

- useR! Conference (July 2027): I am planning to present the work on my data object and functionality package at the conference. It will help me get some feedback as well as network with other researchers on R tools.
- ANZIAM Conference (February 2028): I am planning to present the work on my Project 2, which is about the dynamics of infectious disease modelling, to receive feedback and networking with other researchers in the infectious disease discipline.

Other activities

- I have been co-hosting the Hacky Hour session with Janith over the last year. The Hacky Hour is a weekly session to share and foster knowledge on research software development. I plan to keep running the Hacky Hour session this year as well.
- I was a Volunteer lead on the WOMBAT2025, helping with anything needed on the workshop day.
- I taught a tutorial/workshop with Jayani and Janith on reproducible research in WOMBAT2025 and as part of the NUMBAT workshops. The tutorial introduces Quarto for creating reproducible reporting, academic documents and presentations.

- Tutoring two units, ETC5512: Wild caught data and ETC5521: Diving deeply into data exploration.
- Complete the two PhD compulsory units successfully, ETX6500: Statistical inference and ETX6510: Foundations of econometrics.

Acknowledgements

The materials to generate this report are in the following [repository](#). I am committed to using GitHub version control and Quarto (with R) for reproducible research. It should be noted that the full data is private, which means that only those with access to the data can reproduce this report, at present. This research has been approved by the Alfred Health Ethics Committee, Local reference number 591/18.

I also acknowledge the use of ChatGPT (<https://chat.openai.com/>) and Grammarly to improve the text, grammar, and spelling.

References

- Amaral, André Victor Ribeiro, Jonatan A González, and Paula Moraga. 2023. “Spatio-Temporal Modeling of Infectious Diseases by Integrating Compartment and Point Process Models.” *Stochastic Environmental Research and Risk Assessment* 37 (4): 1519–33.
- Ben-Eliezer, Omri, Talya Eden, Joel Oren, and Dimitris Fotakis. 2022. “Sampling Multiple Nodes in Large Networks: Beyond Random Walks.” In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 37–47.
- Botan, Vanessa, Graham Law, Despina Laparidou, Viet-Hai Phung, Ffion Curtis, Gregory Whitley, Joseph Akanuwe, et al. 2023. “PP63 Variations in the Number of Ambulance Attendances to Care Homes Before and During Covid-19 Pandemic: An Interrupted Time Series Analysis.” BMJ Publishing Group Ltd; the British Association for Accident
- Butts, Carter T. 2008. “Network: A Package for Managing Relational Data in r.” *Journal of Statistical Software* 24 (2). <https://doi.org/10.18637/jss.v024.i02>.
- Cardenas, Nicolas Cespedes, Kimberly VanderWaal, Flávio Pereira Veloso, Jason Onell Ardila Galvis, Marcos Amaku, and Jose HH Grisi-Filho. 2021. “Spatio-Temporal Network Analysis of Pig Trade to Inform the Design of Risk-Based Disease Surveillance.” *Preventive Veterinary Medicine* 189: 105314.
- Chuong Nguyen, Quoc. 2025. “Network Sampling: An Overview and Comparative Analysis.” *arXiv e-Prints*, arXiv–2504.
- Cranmer, Skyler J, Philip Leifeld, Scott D McClurg, and Meredith Rolfe. 2017. “Navigating the Range of Statistical Tools for Inferential Network Analysis.” *American Journal of Political Science* 61 (1): 237–51.

- Csárdi, Gábor, Tamás Nepusz, Vincent Traag, Szabolcs Horvát, Fabio Zanini, Daniel Noom, and Kirill Müller. 2026. *igraph: Network Analysis and Visualization in r*. <https://doi.org/10.5281/zenodo.7682609>.
- Fernández-Gracia, Juan, Jukka-Pekka Onnela, Michael L Barnett, Víctor M Eguíluz, and Nicholas A Christakis. 2017. “Influence of a Patient Transfer Network of US Inpatient Facilities on the Incidence of Nosocomial Infections.” *Scientific Reports* 7 (1): 2930.
- Gohel, David, and Panagiotis Skintzos. 2025. *Ggiraph: Make 'Ggplot2' Graphics Interactive*. <https://doi.org/10.32614/CRAN.package.ggiraph>.
- Gruber, Isabella, Ursel Heudorf, Guido Werner, Yvonne Pfeifer, Can Imirzalioglu, Hanns Ackermann, Christian Brandt, Silke Besier, and Thomas A Wichelhaus. 2013. “Multidrug-Resistant Bacteria in Geriatric Clinics, Nursing Homes, and Ambulant Care—Prevalence and Risk Factors.” *International Journal of Medical Microbiology* 303 (8): 405–9.
- Harmsen, AMK, Georgios F Giannakopoulos, Patrick R Moerbeek, Elise P Jansma, HJ Bonjer, and Frank W Bloemers. 2015. “The Influence of Prehospital Time on Trauma Patients Outcome: A Systematic Review.” *Injury* 46 (4): 602–9.
- Harris, Anthony, and Anurag Sharma. 2018. “Estimating the Future Health and Aged Care Expenditure in Australia with Changes in Morbidity.” *PloS One* 13 (8): e0201697.
- Hoff, Peter D, Adrian E Raftery, and Mark S Handcock. 2002. “Latent Space Approaches to Social Network Analysis.” *Journal of the American Statistical Association* 97 (460): 1090–98.
- Hu, Pili, and Wing Cheong Lau. 2013. “A Survey and Taxonomy of Graph Sampling.” *arXiv Preprint arXiv:1308.5865*.
- Jiao, Bo. 2024. “Sampling Unknown Large Networks Restricted by Low Sampling Rates.” *Scientific Reports* 14 (1): 13340.
- John Jr, Joseph F, and Bruce S Ribner. 1991. “Antibiotic Resistance in Long-Term Care Facilities.” *Infection Control & Hospital Epidemiology* 12 (4): 245–50.
- Kamada, Tomihisa, and Satoru Kawai. 1989. “An Algorithm for Drawing General Undirected Graphs.” *Information Processing Letters* 31 (1): 7–15. [https://doi.org/https://doi.org/10.1016/0020-0190\(89\)90102-6](https://doi.org/https://doi.org/10.1016/0020-0190(89)90102-6).
- Kearney, Anne R, and Daniel Winterbottom. 2006. “Nearby Nature and Long-Term Care Facility Residents: Benefits and Design Recommendations.” *Journal of Housing for the Elderly* 19 (3-4): 7–28.
- Krivitsky, Pavel N, and Mark S Handcock. 2008. “Fitting Latent Cluster Models for Networks with Latentnet.” *Journal of Statistical Software* 24: 1–23.
- Krivitsky, Pavel N, Mark S Handcock, Adrian E Raftery, and Peter D Hoff. 2009. “Representing Degree Distributions, Clustering, and Homophily in Social Networks with Latent Cluster Random Effects Models.” *Social Networks* 31 (3): 204–13.
- Krivitsky, Pavel N., David R. Hunter, Martina Morris, and Chad Klumb. 2023. “ergm 4: New Features for Analyzing Exponential-Family Random Graph Models.” *Journal of Statistical Software* 105 (6): 1–44. <https://doi.org/10.18637/jss.v105.i06>.
- Leifeld, Philip, Skyler J Cranmer, and Bruce A Desmarais. 2018. “Temporal Exponential Random Graph Models with Btergm: Estimation and Bootstrap Confidence Intervals.” *Journal of Statistical Software* 83: 1–36.

- Müller, Kirill, and Hadley Wickham. 2025. *Tibble: Simple Data Frames*. <https://doi.org/10.32614/CRAN.package.tibble>.
- Nair, Shruti Premshankar, Ashley L Quigley, Aye Moa, Abrar Ahmad Chughtai, and Chandini Raina Macintyre. 2023. “Monitoring the Burden of COVID-19 and Impact of Hospital Transfer Policies on Australian Aged-Care Residents in Residential Aged-Care Facilities in 2020.” *BMC Geriatrics* 23 (1): 507.
- Parohan, Mohammad, Sajad Yaghoubi, Asal Seraji, Mohammad Hassan Javanbakht, Payam Sarraf, and Mahmoud Djalali. 2020. “Risk Factors for Mortality in Patients with Coronavirus Disease 2019 (COVID-19) Infection: A Systematic Review and Meta-Analysis of Observational Studies.” *The Aging Male* 23 (5): 1416–24.
- Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- Pedersen, Thomas Lin. 2024a. *Ggraph: An Implementation of Grammar of Graphics for Graphs and Networks*. <https://doi.org/10.32614/CRAN.package.ggraph>.
- . 2024b. *Tidygraph: A Tidy API for Graph Manipulation*. <https://CRAN.R-project.org/package=tidygraph>.
- Rao, K Venkateswara, A Govardhan, and KV Chalapati Rao. 2012. “Spatiotemporal Data Mining: Issues, Tasks and Applications.” *International Journal of Computer Science and Engineering Survey* 3 (1): 39.
- Robins, Garry, Pip Pattison, Yuval Kalish, and Dean Lusher. 2007. “An Introduction to Exponential Random Graph (p^*) Models for Social Networks.” *Social Networks* 29 (2): 173–91.
- Silk, Matthew J, Darren P Croft, Richard J Delahay, David J Hodgson, Nicola Weber, Mike Boots, and Robbie A McDonald. 2017. “The Application of Statistical Network Models in Disease Research.” *Methods in Ecology and Evolution* 8 (9): 1026–41.
- Strausbaugh, Larry J, Kent B Crossley, Brenda A Nurse, Lauri D Thrupp, SHEA Long-Term-Care Committee, et al. 1996. “Antimicrobial Resistance in Long-Term-Care Facilities.” *Infection Control & Hospital Epidemiology* 17 (2): 129–40.
- Tandan, Meera, Poshan Thapa, Preeti Maharjan, and Buna Bhandari. 2022. “Impact of Antimicrobial Stewardship Program on Antimicrobial-Resistance and Prescribing in Nursing Homes: A Systematic Review and Meta-Analysis.” *Journal of Global Antimicrobial Resistance* 29: 74–87.
- van der Meer, Lucas, Lorena Abad, Andrea Gilardi, and Robin Lovelace. 2024. *Sfnetworks: Tidy Geospatial Networks*. <https://luukvdmeer.github.io/sfnetworks/>.
- Wang, Earo, Dianne Cook, and Rob J Hyndman. 2020. “A New Tidy Data Structure to Support Exploration and Modeling of Temporal Data.” *Journal of Computational and Graphical Statistics* 29 (3): 466–78. <https://doi.org/10.1080/10618600.2019.1695624>.
- Wickham, Hadley. 2014. “Tidy Data.” *Journal of Statistical Software* 59 (10): 1–23. <https://doi.org/10.18637/jss.v059.i10>.
- . 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://doi.org/10.32614/CRAN.package>.

`dplyr`.

- Wyer, Leanna, Yair Guterman, Vivian Ewa, Eddy Lang, Peter Faris, and Jayna Holroyd-Leduc. 2024. “The Impact of the COVID-19 Pandemic on Transfers Between Long-Term Care and Emergency Departments Across Alberta.” *BMC Emergency Medicine* 24 (1): 9.
- Yutani, Hiroaki. 2025. *Gghighlight: Highlight Lines and Points in 'Ggplot2'*. <https://CRAN.R-project.org/package=gghighlight>.
- Zhou, Zhengyi, David S Matteson, Dawn B Woodard, Shane G Henderson, and Athanasios C Micheas. 2015. “A Spatio-Temporal Point Process Model for Ambulance Demand.” *Journal of the American Statistical Association* 110 (509): 6–15.

Listing 4 Code for creating an interactive network visualisation. Instead of using the ggraph function, such as geom_node_point() and geom_edge_line, geom_point_interactive and geom_segment_interactive are used. The implementation is more complex than necessary because the layout information is not easily accessible.

```
interactive_vis <- simple_graph |>
  mutate(name = str_remove(name, "'")) |>
  ggraph(x = long, y = lat) +
  geom_sf(data = vic_map, color = "white") +
  geom_segment_interactive(
    data = simple_graph |>
      activate(edges) |>
      as_tibble() |>
      mutate(id = row_number()),
    alpha = 0.2,
    aes(x = long_racf,
        y = lat_racf,
        xend = long_hosp,
        yend = lat_hosp,
        tooltip = weight,
        data_id = id)) +
  geom_point_interactive(aes(x = x,
                             y = y,
                             color = category,
                             tooltip = name,
                             data_id = name)) +
  scale_color_brewer(name = "Facility", palette = "Set1")

girafe(ggobj = interactive_vis,
       options = list(
         opts_hover(
           css = "fill:lightblue;stroke:grey;stroke-width:0.5px"
         ),
         opts_zoom(min = 0.5, max = 3)
       ))
```