Exploratory analysis of time-use activity data using network theory

Eoghan McKenna[1], Sarah Higginson[2], Tom Hargreaves[3], Jason Chilvers[3] and Murray Thomson[1]

[1]Centre for Renewable Energy Systems Technology (CREST),
Wolfson School of Mechanical, Manufacturing and Electrical Engineering
Loughborough University, LE11 3TU, UK

[2]Environmental Change Institute, School of Geography and the Environment,
Oxford University Centre for the Environment, University of Oxford, South Parks Road, Oxford, OX1 3QY

[3]Science, Society and Sustainability (3S) Research Group,
School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

Correspondence: e.j.mckenna@lboro.ac.uk

Paper prepared for DEMAND Centre Conference, Lancaster, 13-15 April 2016

Abstract

National time-use surveys provide detailed data of people's activities over a 24-hour period and in numerous ways have proven to be useful sources of information for energy demand research. Crucially, such data is used by both engineers and social scientists in energy demand research, and as such it can serve as a useful interdisciplinary 'bridge' within this field of research. This paper presents a novel method for the visualisation and analysis of time-use activity data which is critically distinguished from conventional methods by explicitly representing how activities are dynamically interconnected in time. The method uses network theory as a formal framework for diagramming sequences of activities to create dynamic network graphs. The method provides insight by visually revealing how activities can be classified into a range of different types depending on their position and role within the overall activity network. The network metrics of degree and centrality are used to identify personal care (e.g. eating, washing and dressing), household and family care, and travel as the categories of activities with the highest degree and centrality, suggesting that they act as 'anchors' and 'hubs' within the overall activity network. These types of metrics are used when assessing other types of networks such as information or technical networks and can be useful for identifying what make them flexible or resilient to change. This work provides an initial step towards the goal of providing similarly beneficial analysis of activities, with the view of informing research into what makes people's activities, and by extension their energy demand, flexible or resilient to change.

Keywords: Time-use; activity; network theory; network analysis; flexible demand; energy demand.

1. Introduction

Studies that explore potential pathways to securing a low-carbon future for the UK are characterised by the considerable changes required to the way electricity is generated, distributed and used (Foxon, 2013; Barton et al., 2013; Barnacle et al., 2013). The large-scale deployment of low-carbon generation technologies such as renewable energy, nuclear and carbon capture and storage poses a particular challenge associated with the task of balancing electricity supply and demand, and the prospect of potentially considerable amounts of surplus generation, unmet demand, and flexible thermal generation operating at uneconomically low capacity factors means there is a strong requirement for measures such as interconnection, energy storage and flexible demand to help balance the system (Delucchi and Jacobson, 2011; Elliston et al., 2012; Budischak et al., 2013; Rasmussen et al., 2012). It is important therefore to evaluate and compare the impact these measures can have on securing a low-carbon power system, for example through techno-economic modelling (D. Pudjianto et al., 2014). These studies provide useful assessments of the 'whole-system' value and role that each measure can play, however it is important that the modelling assumptions are well-founded. Flexible demand is an area where there is still considerable uncertainty and this is a key area where there is a need for interdisciplinary research to understand the potential for flexible demand, its impact on people's everyday lives, and to improve engineering models with assumptions that are feasible, realistic and that are founded on sound social science.

With this in mind, this paper focuses on activities as an interdisciplinary 'bridge' between the social sciences and engineering models, as data about people's activities is used by both engineers and social scientists in energy demand research. For example, social practice theorists use time-use activity data as a useful basis for inferring insights about practice flexibility indirectly (Shove et al., 2009), while engineers use the same data as a basis for constructing 'bottom-up' models to simulate end-use energy demand within the built-environment (Widén et al., 2012; Keirstead and Sivakumar, 2012). The representation of activities within energy models is, however, rather crude, for example by assuming that the probability of being in any activity state at a time period is only dependent on the activity state in the previous time period. While modelling techniques that employ such simplifications can nonetheless capture much of the statistical diversity of people's activities in the real world (McKenna et al., 2015), there is clearly a benefit to improving them, especially where the desire is to incorporate complexities such as flexible demand within the models.

2. Energy demand flexibility and the connections between activities

People's lives are full of practices and they are all interconnected (Shove et al., 2012; Watson, 2012). Moving one in time cannot be done without impacting others. If we're interested in shifting one practice, we need to understand how it is connected to others. We might be interested in which practice is the most connected, as we might hypothesise that highly connected practices are the most inflexible or difficult to shift. Indeed, if we are interested in making laundry more flexible, for instance, it may be just as important to focus our efforts on shifting the practices that laundry is most connected to (e.g. clothes wearing/fashion, school uniform rules etc.) as on the practice of laundry itself (Higginson et al., 2013). The temporal sequencing of practices is therefore of particularly importance when considering practice flexibility. Parallel work by the authors has focussed on the connections within practices (Higginson et al., 2015) while here the focus is on the connections between activities[1].

There are many ways in which the connections between activities can be considered, but here the focus will be on how they are connected in time through their temporal sequencing. As in the authors' previous work, network theory will be used as a method for visualising and analysing such sequences. Network theory is a broad field of research which is devoted to the scientific analysis of the structure and dynamics of networks including technical, information, biological and social networks (Newman, 2003; Strogatz, 2001; Watts and Strogatz, 1998). Of particular relevance for the purposes of the present work, it has proven to be a useful method for understanding networks that consist of flows, traffic, or sequences, such as the diffusion of viruses or innovations through social networks (Watts, 2002).

[1] We are not implying that practices are analogous to activities, but they can nonetheless provide a useful basis for inferring insights about practice flexibility indirectly, for example as shown in (Shove et al., 2009). The distinction between activities and practices is complex (Schatzki, 1996), but for the purposes of clarity we offer the simple definition that activities are descriptions of what people do, while practices also involve why, how, where, when and with whom and what.

Figure 1 shows an example of a sequence of activities for one person. The nodes are activities, and the connections show their sequence or direction in time. The direction or 'flow' of the graph is from left to right in this case, though there are 'feedback loops' when the person performs an activity multiple times. The person gets dressed, then eats, travels to work, breaks for lunch, starts work again, travels home, then does the laundry. Though this is a trivial example, the usefulness of this type of graph is that it can reveal connections between activities, and when scaled up can highlight particularly important sequences of activities – which are likely to be critical in attempting to understand patterns of energy demand – or whether certain activities may act as 'hubs', in the same sense that websites such as Amazon or Wikipedia act as the hubs in the world wide web.

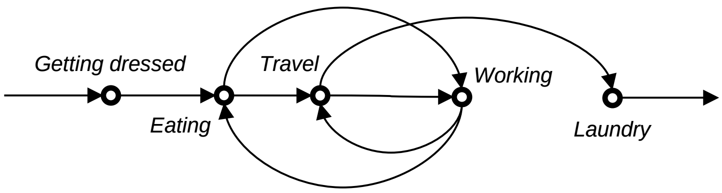

Figure 1 - a graph of practices (nodes) and their sequential flow in time (arrows).

3. Analysis of time-use activity data using network theory

In this section network theory is used to analyse the UK time-use survey data (Ipsos-RSL and Office for National Statistics, 2003). The time-use survey data consists of 24-hour diaries detailing people's activities, at 10-minute resolution. The UK time-use survey data consists of 20,981 diaries. A sub-set of 500 diaries are used here to illustrate the analysis technique. Figure 2 provides a specific example of the type of graph shown in Figure 1 but based on actual activity data. It consists of a single time-use diary for a single person, and shows the individual activities the person performed during their day, and connects these together using arrows to illustrate the sequence or flow of activities. Some activities have 'self-loops' – lines that start and end at the same node – which indicate where activities persisted from one time period to the next. The activities (nodes) are laid out such that those that are connected are closer together, which illustrates which activities are closer together in terms of their sequencing, potentially revealing clusters of activities. Figure 2 for example shows that for this person various travel activities are closely connected. The nodes are coloured according to the activity category, as defined in the time-use survey. The nodes have been sized according to the centrality of the node, which is a measure of the importance of a node within a network, and will be described in more detail below. Eating can be seen to be the most central activity in this person's sequence of activities.

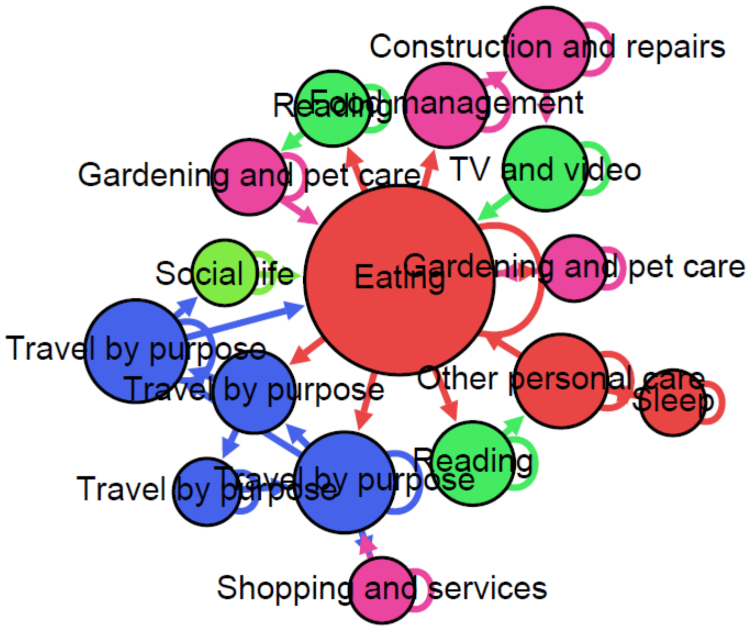

Figure 2 – example network graph of a single time-use activity diary entry.

Different people have different sequences of activities, and their network graphs reflect this. Figure 3, for example, provides a contrasting example. There is no dominantly central activity in this network. Personal care activities such as eating and sleeping appear as a sequence on one side of the graph, while social activities are clustered together on another side. Media activities, in this case reading and watching the TV are not grouped together but are interspersed between other types of activity.

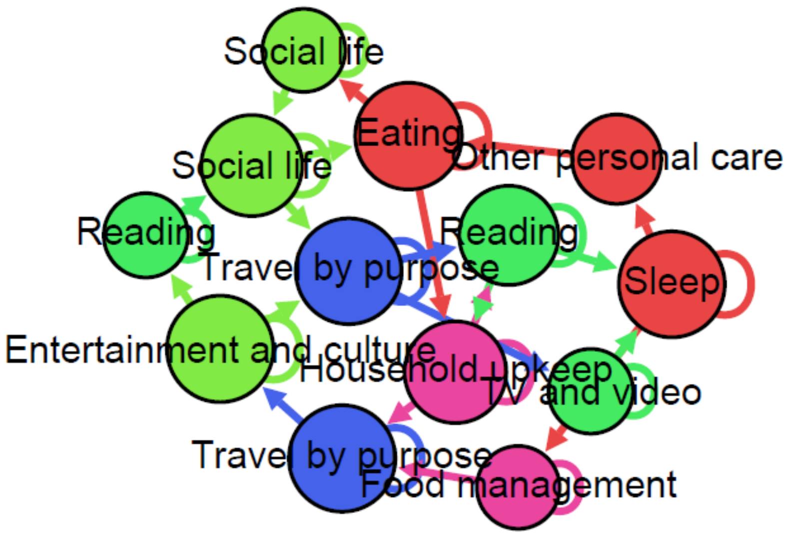

Figure 3 – contrasting example of a network graph of another time-use activity survey diary.

Analysis of individual networks such as these is useful to introducing this new way of structuring this type of data and introducing some of the related concepts. One of the strengths of the technique however is that it is readily scaled up to analyse large numbers of activity graphs. This is important for the approach to be able to provide insight into society-wide patterns of practice and energy demand and for revealing ‘macro’ features and relationships in data that might be too large or complex to uncover simply by inspection of the raw data alone. As a proof of concept therefore the following will focus on the analysis of a 500 individual network graphs, which are shown in Figure 4 for illustrative purposes. Each cluster is an individual activity network associated with one of the diary entries, while the layout is random.

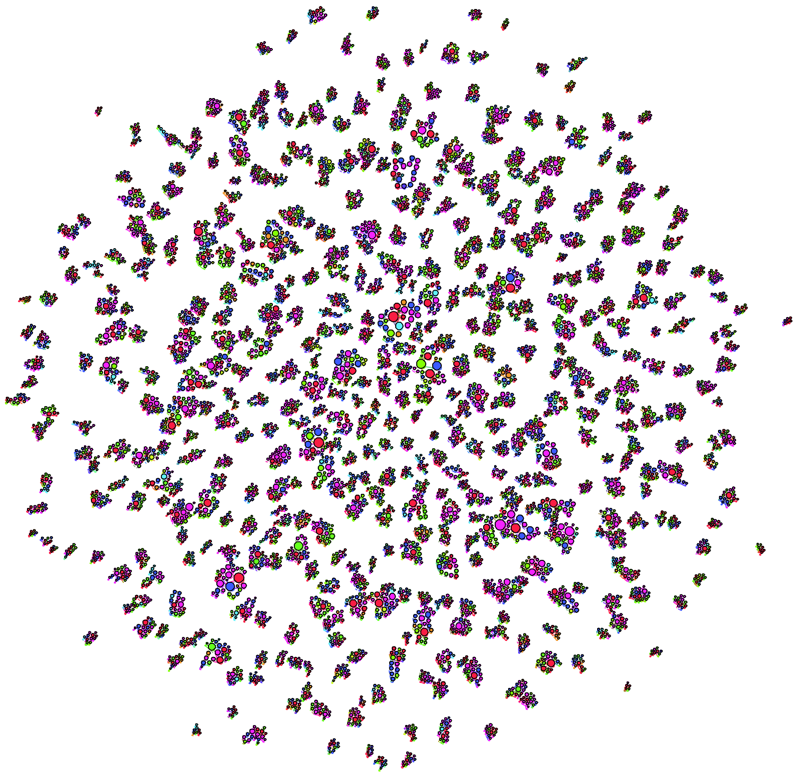

Figure 4 – illustrative graph showing 500 individual activity network graphs.

The first metric to be analysed is the degree of the activities, which is a measure of how many activities a given activity is connected to. The degree is an interesting metric to consider as it can be hypothesised that activities with high degree act as 'anchors' in people's lives as they are more connected to other activities e.g. the eating activity in Figure 2. Anchors could therefore be important activities to focus on to understand the structure and dynamics of activities, and ultimately what makes people's lives flexible or resilient to change. We can hypothesise that anchors might be particularly effective activities to try and change, as they are likely to have a big overall impact on the rest of the network. At the same time, it may well be that precisely because of their central importance they are also particularly difficult to shift.

Figure 5 shows the distribution of degrees for every individual activity node shown in the previous figure. The distribution appears to follow a Poisson distribution, with the majority of activities connected to between 2 and 6 activities, and an average of 4.8. The distribution however has a 'long tail' with a small number of activities achieving a very high degree – up to 17.

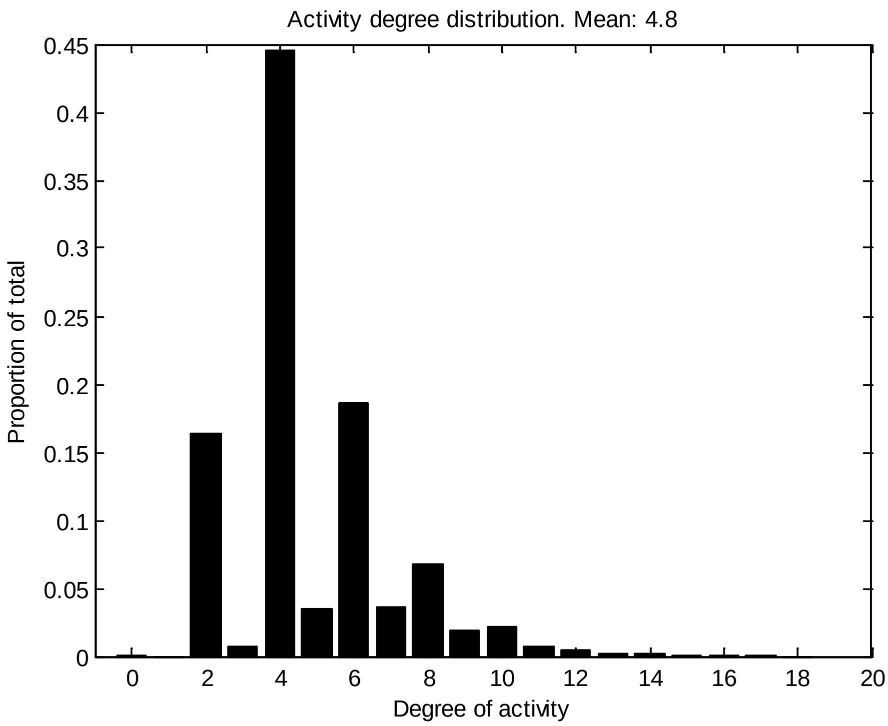

Figure 5 – Degree distribution for all activities. Average degree is 4.8.

Figure 6 shows how the degree distribution varies according to the sub-category of activity, as recorded in the time-use survey. The figure shows box-plots for each type of sub-category. The median is shown by a circled black square, a blue box extends to the 25th and 75th percentiles, whiskers extend to the most extreme points of the distribution, and outliers are plotted individually. The results show that the majority of activities within each sub-category fall within the range 4-6 degree. There are some with significantly higher degree, however, in particular eating, other personal care (which includes washing and dressing), and to a lesser extent childcare and caring for other adult members of the household. The relatively high degree of activities within these sub-categories would indicate that these activities tend to be anchors within people's everyday lives.

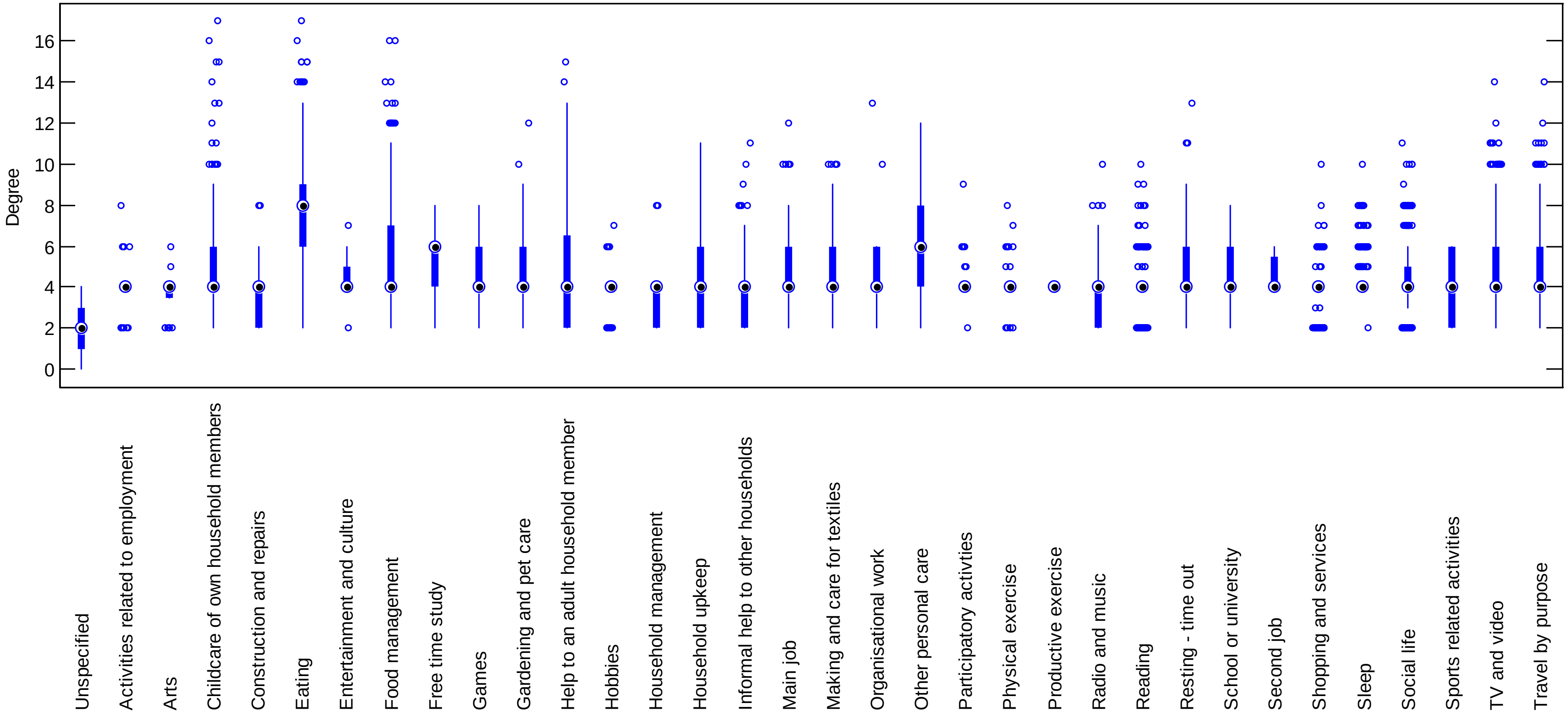

Figure 6 – degree distribution broken down by activity sub-category according to the time-use survey.

Degree is a measure of the number of direct one-to-one connections between activities. Activities are however part of longer sequences than this and 'centrality' can be a useful network metric to consider as it provides a measure of the importance of a node within the wider network. A node has a high centrality if it appears on many 'shortest paths' between other nodes. For information and technical networks such the World Wide Web, or the Internet, this provides a useful measure of the amount of traffic the node can expect to receive. While the activity networks considered here are not directly analogous to flow networks such as the Internet, centrality nonetheless gives a way of quantifying the importance of activities within the wider network and a way of identifying activities that act as central 'hubs' within people's everyday lives.

Figure 7 shows the distribution of centrality for each individual activity node within the 500 diary graphs shown previously. The distribution has been plotted on a log-log scale to reveal that there appears to be a truncated power law relationship between centrality and the number of activities with that centrality. What this means is that there are a very large number of activities that have low centrality, i.e. they are peripheral to the network, while there are a very few activities with a disproportionally large centrality, and that therefore play a correspondingly larger role within the network as a result. The presence of a power law in a network property is characteristic of many real-world networks (Strogatz, 2001)– for example the World Wide Web, which has a few very highly connected nodes (e.g. Yahoo, Amazon, Wikipedia etc.) and a very large number of sites with very few connections. Such networks where power laws arise are called 'scale-free' as no single characteristic scale can be defined for the network. It means that, from a network centrality perspective, activity networks are similar to scale-free networks such as the internet, and that they can be similarly dominated by the equivalent of the Amazon's of the world of activity.

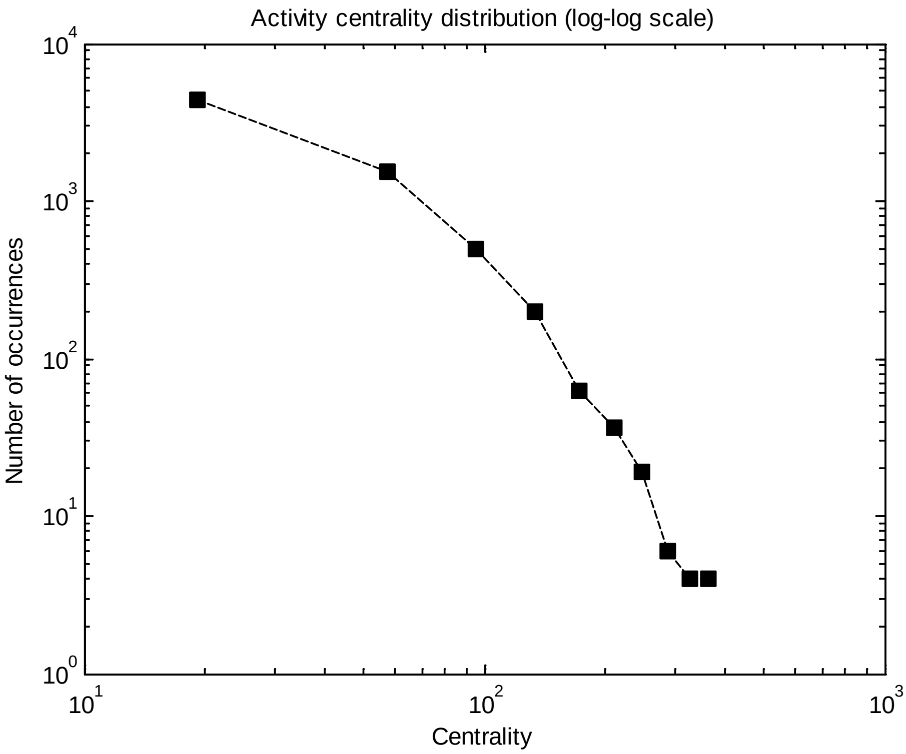

Figure 7 – Centrality distribution for all activities.

Figure 8 shows the relationship between an activity's degree and its centrality, which confirms the intuitive result that activities with higher degree are generally those with higher centrality. The results also show however that some activities with low degree can nonetheless be highly central, and some with high degree can have low centrality.

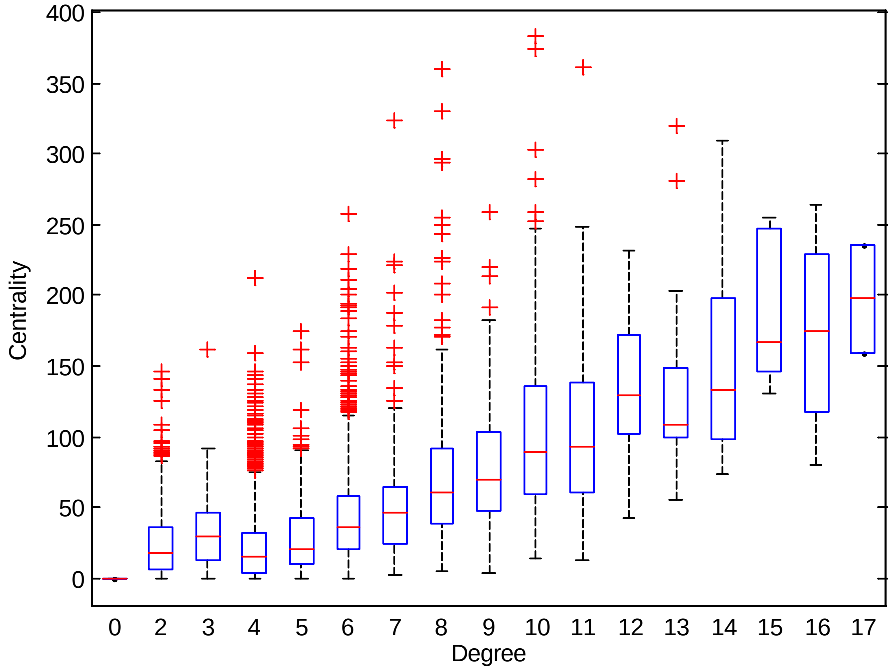

Figure 8 – relationship between activity degree and centrality.

Figure 9 shows the results of grouping the activities according to their category, as specified in the time-use survey, and illustrates whether there are general differences in centrality according to the category of the activity in question. Boxplots are used as in Figure 6 above. The results show that personal care (which includes eating, washing and dressing, and sleeping), household and family care, and travelling, are generally the most central activities within people's sequences of activities, and so would tend to act as the main hubs through which the sequence of people's activities tends to flow. Sports, study and employment have the lowest levels of centrality.

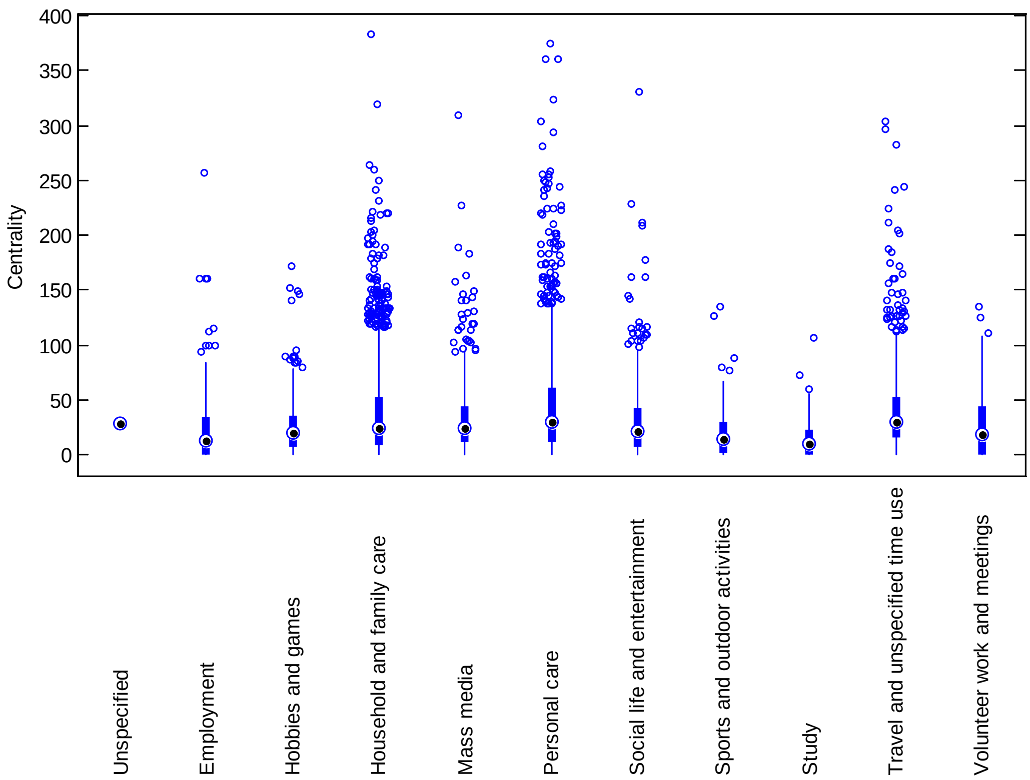

Figure 9 – boxplots showing distribution of centrality for activities grouped by the categories used in the time-use survey.

The categories used in the time-use survey are quite broad and can often include many individual types of activity. Figure 10 therefore shows the mean centrality for every type of activity reported in the 500 diaries i.e. they are not grouped according to their category or sub-category. The distribution appears to have two main regions, with an initial region with steeply descending centrality, which levels off after approximately the first 20 activities to a slow descent over the remaining ~180 activities.

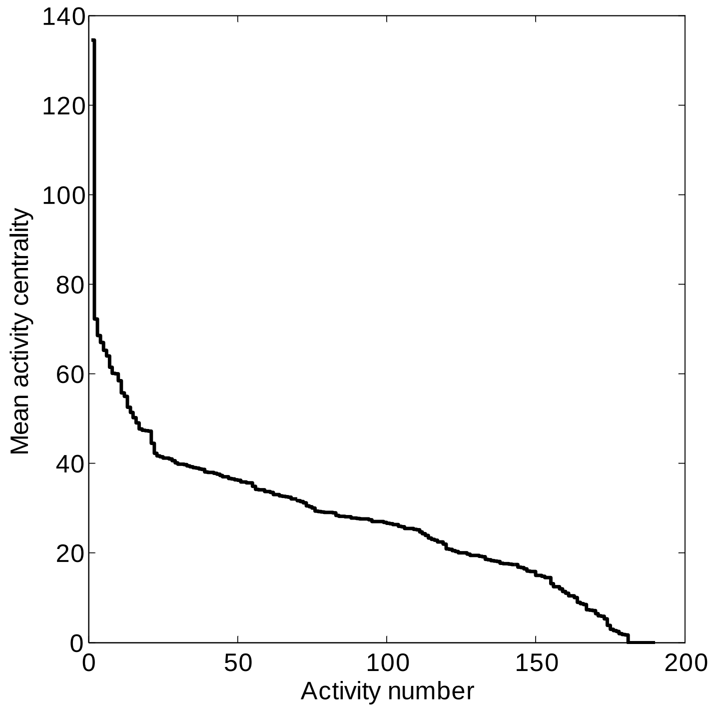

Figure 10 – ranked distribution of average centrality for each activity type.

Table 1 lists the top 20 activities, which roughly correspond to those that appear in the first region of the distribution above. Some of the entries such as 'other specified water sports' or 'disposal of waste' appear because they are highly central activities in a small number of diaries. For example, Figure 11 shows the only diary entry in which 'other specified water sports', and Figure 12 shows one of only four diary entries in which 'disposal of waste' appears. The point is that while these activities played central roles within these particular people's daily activities, we cannot make inferences about these activities within the broader population, as we can for activities such as eating, travel, and household and family care which the results again emphasise are, in general, highly central activities within the networks of people's activities.

Table 1 – top ten types of activities according to mean centrality across the 500 graphs.

Activity label	Mean centrality
'Other specified water sports'	135
'Eating'	72
'Travel escorting to/ from education'	69
'Visiting a botanical site'	67
'Other specified physical care & supervision of a child'	65
'Unspecified household upkeep'	64
'Disposal of waste'	62
'Travel related to household care'	60
'Heating and water'	60
'Food preparation'	58
'Physical care & supervision of an adult household member'	56
'Travel related to hobbies other than gambling'	55
'Construction and repairs as help'	53
'Information searching on the internet'	51
'Travel related to shopping'	50
'Food management as help'	49
'Travel related to religious activities'	48
'Wash and dress'	47
'Illegible activity'	47
'Travel escorting a child (other than education)'	47

Figure 11 - single activity diary where water sports has high centrality.

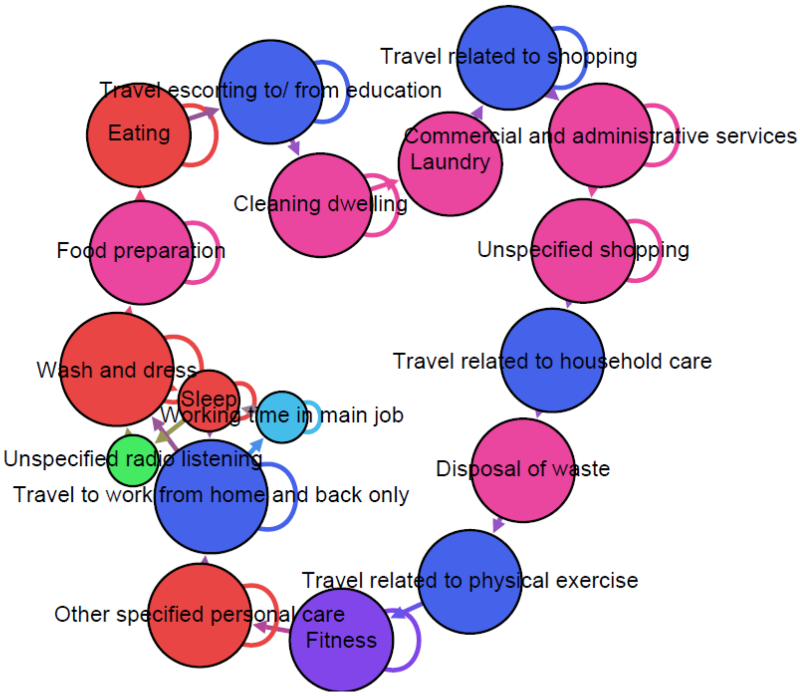

Figure 12 – single diary entry where disposal of waste has a high centrality (relative to all other diaries).

The results have highlighted three categories of activities in particular that, from a network perspective, are particularly important. These are personal care, household and family care, and travel. It is interesting to note that while personal care and household and family care were highlighted as both 'anchors' (degree) and 'hubs' (centrality), travel appears mainly as a hub, and not an anchor. This would align with the idea of travel as a means to enable or support other activities, and not an activity to be performed for its own sake. Furthermore, we note that personal care activities such as preparing food, cooking, washing up, and washing and dressing, as well as travelling activities have relatively high greenhouse gas intensities (Torriti et al., 2015) indicating that the activities that are most central from a network perspective are also those that are amongst the most important to decarbonise.

4. Conclusions

This paper has described the application of network theory to the analysis of a sub-set of the UK time-use survey data. The approach provides a new way of visualising and analysing this type of data that is critically distinguished from conventional methods by explicitly capturing the connections between activities and the structure of the activity network that results. The results have shown that network theory provides a means of differentiating activities based on their level of interconnection with adjacent neighbours (their degree) and the level of their importance when interpreted as conduits through which people's everyday lives flow (their centrality). This allows the identification of activities which, from a network perspective, can be considered as anchors – particularly difficult to shift but likely to have a big overall impact – and hubs – important for maintaining the flow of activities but not necessarily an activity performed for its own sake.

These types of metrics are used when assessing other types of networks such as information or technical networks and can be useful for identifying properties that make them flexible or resilient to change (e.g. to the removal of nodes). Ultimately the aim of this approach is to be able to provide similarly beneficial analysis of activities, with the view of informing research into what makes people's activities, and by extension their energy demand, flexible or resilient to change. This is an ambitious aim, and while we are still a long way from being able to do that, this work nonetheless demonstrates that network theory can be applied to this type of data and can provide new ways of visualising and analysing it.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council, UK, within the Realising Transition Pathways project (EP/K005316/1).

References

Barnacle, M., Robertson, E., Galloway, S., Barton, J., Ault, G., 2013, Modelling generation and infrastructure requirements for transition pathways. Energy Policy, 52, 60-75.

Barton, J., Huang, S., Infield, D., Leach, M., Ogunkunle, D., Torriti, J., et al., 2013, The evolution of electricity demand and the role for demand side participation, in buildings and transport. Energy Policy, 52, 85-102.

Budischak, C., Sewell, D., Thomson, H., Mach, L., Veron, D.E., Kempton, W., 2013, Cost-minimized combinations of wind power, solar power and electrochemical storage, powering the grid up to 99.9% of the time. Journal of Power Sources, 225, 60-74.

D. Pudjianto, M. Aunedi, P. Djapic, G. Strbac, 2014, Whole-Systems Assessment of the Value of Energy Storage in Low-Carbon Electricity Systems. IEEE Transactions on Smart Grid, 5, 1098-1109.

Delucchi, M.A., Jacobson, M.Z., 2011, Providing all global energy with wind, water, and solar power, Part II: Reliability, system and transmission costs, and policies. Energy Policy, 39, 1170-1190.

Elliston, B., Diesendorf, M., MacGill, I., 2012, Simulations of scenarios with 100% renewable electricity in the Australian National Electricity Market. Energy Policy, 45, 606-613.

Foxon, T.J., 2013, Transition pathways for a UK low carbon electricity future. Energy Policy, 52, 10-24.

Higginson, S., McKenna, E., Hargreaves, T., Chilvers, J., Thomson, M., 2015, Diagramming social practice theory: An interdisciplinary experiment exploring practices as networks. Indoor and Built Environment, 24, 950-969.

Higginson, S., Thomson, M., Bhamra, T., 2013, “For the times they are a-changin”: the impact of shifting energy-use practices in time and space. Local Environment, 19, 520-538.

Ipsos-RSL and Office for National Statistics, 2003., United Kingdom Time Use Survey, 2000 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], September 2003. SN: 4504. [Available online: http://discover.ukdataservice.ac.uk/catalogue?sn=4504].

Keirstead, J., Sivakumar, A., 2012, Using Activity-Based Modeling to Simulate Urban Resource Demands at High Spatial and Temporal Resolutions. Journal of Industrial Ecology, 16, 889-900.

McKenna, E., Krawczynski, M., Thomson, M., 2015, Four-state domestic building occupancy model for energy demand simulations. Energy and Buildings, 96, 30-39.

Newman, M.E., 2003, The structure and function of complex networks. SIAM Review, 45, 167-256.

Rasmussen, M.G., Andresen, G.B., Greiner, M., 2012, Storage and balancing synergies in a fully or highly renewable pan-European power system. Energy Policy, 51, 642-651.

Schatzki, T.R., 1996. Social practices: A Wittgensteinian approach to human activity and the social, 1st ed. Cambridge University Press, Cambridge.

Shove, E., Pantzar, M., Watson, M., 2012. The dynamics of social practice: everyday life and how it changes, 1st ed. Sage, Thousand Oaks, CA.

Shove, E., Trentmann, F., Wilk, R., 2009. Time, consumption and everyday life: practice, materiality and culture, Berg,.

Strogatz, S.H., 2001, Exploring complex networks. Nature, 410, 268-276.

Torriti, J., Hanna, R., Anderson, B., Yeboah, G., Druckman, A., 2015, Peak residential electricity demand and social practices: Deriving flexibility and greenhouse gas intensities from time use and locational data. Indoor and Built Environment, 24, 891-912.

Watson, M., 2012, How theories of practice can inform transition to a decarbonised transport system. Journal of Transport Geography, 24, 488-496.

Watts, D.J., Strogatz, S.H., 1998, Collective dynamics of 'small-world'networks. Nature, 393, 440-442.

Watts, D.J., 2002, A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99, 5766-5771.

Widén, J., Molin, A., Ellegård, K., 2012, Models of domestic occupancy, activities and energy use based on time-use data: deterministic and stochastic approaches with application to various building-related simulations. Journal of Building Performance Simulation, 5, 27-44.