The application of network theory in the official statistics

Áron Kincses
Hungarian Central Statistical Office

"I think the next century will be the century of complexity."

Stephen Hawking

The challenges faced by official statistics in the 21st century are manifold. We are surrounded by systems that are becoming substantially more and more complex. With the emergence of new phenomena (e.g. globalisation, global demographic trends or sustainable development) and complex realities that need to be meaningfully and timely captured by official statistics, new patterns, routes and types of data have also been emerging, offering opportunities to improve the relevance of statistics. These trends also offer new, potentially usable tools and methods for the measurement of these changing phenomena.

Network theory is applied in multiple disciplines, including computer science, business, economics, biology, particle physics, operations research or sociology, but it has so far avoided the mainstream of official statistics. Network science is defined not only by its subject matter, but also by its methodology. This methodology offers a common language in which statisticians are not yet fluent.

The social network determines and reveals the spread of knowledge, behaviour and resources. Consider a society that requires cooperation among billions of people, and communications infrastructures that join billions of cell phones with computers and satellites. The communication networks, describing which communication devices interact with each other, are at the centre of the modern communication system. The business network is owned by the business enterprise, where the aim of the network is to support the informational and operational requirements of the business such as marketing, sales, accounting, and manufacturing departments. The variety of relationships within MNEs between parent and affiliations underpin the importance of dynamic capabilities in the global market. Trade networks maintain our ability to exchange goods and services, being responsible for material

well-being. These systems are collectively called complex systems, capturing the fact that it is difficult to derive their collective behaviour from the knowledge of the system's components.

Attempting to provide a short overview of area of use for the network theory, the aim of this paper is to share innovations and some of the most important tangible outcomes of network analysis in official statistics (e.g. social network analysis, business relationship (EGR) and migration settlements network), which helps us better understand globalisation and social networks, reduce the burden of data providers, and achieve more accurate and efficient use of dig data. The other goal is to show how network analysis can be used to help the resourceful operation of statistical offices (namely, planning, controlling and management support). The statistical offices themselves "produce" a lot of big data, but we might think them less important than the data sources that are out of the office. Through examples I shall present how the network analysis can support the modernisation processes of statistical offices: the stakeholder analysis, use of the phone data and the electronic access control system of NSI to reveal the real processes of statistical offices in comparison the theoretical ones.

In conclusion, network theory is a new innovative tool and thinking in our changing word, which can help us handle the challenges and opportunities of the 21st century. The networks do not stop at the borders of the countries, nor can they be examined effectively at country level: they require collaboration at EU level. Given the important role complex systems play in our daily life and in our economy, understanding and eventually controlling them is one of the major intellectual and scientific challenges of the 21st century. It is a challenge that European statistics cannot afford not to tackle.

REFERENCES

1. Albert-László Barabási (2016): Network Science, Cambridge University Press, p.453
2. Albert-László Barabási (2010): Bursts: The Hidden Pattern Behind Everything We Do; Button, New York; p. 310
3. Battiston, Federico; Nicosia, Vincenzo; Latora, Vito (2017): The new challenges of multiplex networks: Measures and models. The European Physical Journal Special Topics. 226 (3): 401–416. doi:10.1140/epjst/e2016-60274-8. ISSN 1951-6355.

4. Cardillo, A.; et al. (2013): Emergence of network features from multiplexity. Scientific Reports. 3:1344.

5. D Shah and T. Zaman (2011): Rumours in a Network: Who's the Culprit? IEEE Trans. Inform. Theory 57:pp. 5163-5187

6. Kryven, Ivan (2016-07-27): Emergence of the giant weak component in directed random graphs with arbitrary degree distributions". Physical Review E. 94 (1): 012315. doi:10.1103/PhysRevE.94.012315.

7. Lawyer, Glenn (March 2015): Understanding the spreading power of all nodes in a network. Scientific Reports. 5 (O8665):8665.

8. Lundy Lewis (2001) Managing Business and Service Networks. p. 138

9. Newman, M.E.J. (2010): Networks: An Introduction. Oxford University Press. ISBN 978-0199206650

10. T. W. Vante (1995): Network models of the diffusion of innovations. Hampton Press, Cresskill, NJ; p.188