EXperimental Learning

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Big Data and Social Analytics certificate course

MODULE 4 UNIT 1
The structure of networks

MIT BDA Module 4 Unit 1 Video Resource

Learning outcomes:

LO3: Interpret the structure of peer networks.

LO4: Articulate how data techniques can be used to reveal peer networks.

Title: The structure of networks

In Video 4, Xiaowen Dong introduces the concept of network centrality measures, which refer to the degree, closeness, betweenness, and eigenvector centrality of a network.

Video 4: Xiaowen Dong – Structure of networks I.
(To download the video, click here.)

Although these centrality measures have already been touched on in the above video, should you wish to learn more about them, you are encouraged to read Section 9.1.5.1 of this resource, which further explains the notion of centrality in the context of Zachary's karate club. The resource focuses particularly on degree and betweenness centrality, which are most evident in the social network structure of the karate club. Zachary's karate club is a widely-used social network graph, made popular by anthropologist Wayne Zachary, who based and published an ethnographic study on a small karate club (1977).

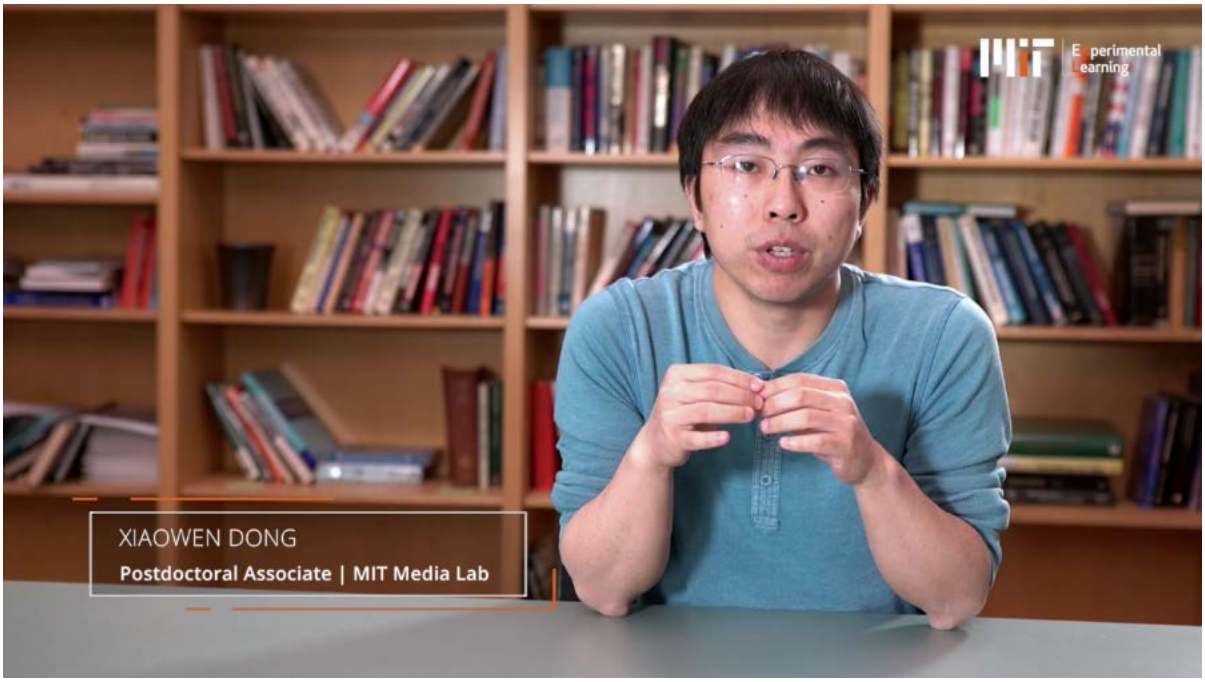
Due to its popularity, academics often use Zachary's karate club to provide context-based explanations of certain concepts related to graph theory, as seen in the resource above. Zachary's karate club, however, cannot be used to explain eigenvector centrality. Further information about this centrality measure can be found in Section VIII C1 of M. E. J. Newman's article titled "The structure and function of complex networks".

If you would like more general information on the centrality measures, refer to the Centralities and Prestige section of the Social Network Visualizer resource.

Note:

All of the articles included in this video resource have been provided for enrichment purposes, and will not form part of the assessments for this module. Additionally, Xiaowen Dong will draw upon the abovementioned karate club example in his explanation of theory covered in the final video of this component.

According to Barabasi and Albert (1999), a common feature of many real-world networks is that they are scale-free, meaning that their degree distributions follow a power-law. In his second structure of networks video (Video 5), Xiaowen Dong introduces the concept of power-law distributions within networks.



Video 5: Xiaowen Dong – Structure of networks II.
(To download the video, click here.)

As part of his explanation in video 5, Xiaowen Dong makes reference to three real-world examples of power-law distributions – actor collaboration, web pages on the internet, and an electric power grid – which originate from research conducted by Barabasi and Albert (1999).

Should you wish to find out more about said research, and gain a more nuanced understanding of power-law distribution beyond the requirements of the assessments on this course, read their article titled "Emergence of Scaling in Random Networks". Additionally, for further information on degree distribution and small-world networks, read Sections III and VI respectively, from M.E.J. Newman's article "The structure and function of complex networks". Lastly, if you have an interest in learning more about small-world and scale-free networks, refer to Sections VI and VII of the paper "Statistical Mechanics of Complex Networks".

Note:

The clustering coefficient introduced in video 5 is called the global clustering coefficient or transitivity. In comparison, Watts and Strogatz (1998) proposed another definition which is referred to as the local clustering coefficient. Details about this definition can be found in the caption of Figure 2 of their paper which is freely available.

In Video 6, Xiaowen Dong introduces and explains the various existing approaches to graph clustering. These approaches are spectral graph partitioning, hierarchical clustering, and the Louvain method.

Video 6: Xiaowen Dong – Graph clustering.
(To download the video, click here.)

For further reading regarding the approaches covered in this video, which will provide you with an understanding beyond what is required for the assessments this course, follow the links referred to below.

It is important to note that the resource on community detection in graphs is quite lengthy. You are advised to only focus on Section IV of the resource (Pages 16-20), which provides digestible

explanations of graph clustering approaches (excluding the Louvain method). For further insight into the background and application of the Louvain method, read the journal article titled “Fast unfolding of communities in large networks”. Additionally, if you would like further information on spectral graph partitioning, the paper titled “A Tutorial on Spectral Clustering” is a notable resource.

You are now ready to apply your knowledge

Now that you’ve worked through all of the videos in the Engage unit, you are ready to apply your newly gained knowledge by completing the activities in the Apply unit. You can access these activities by navigating back to your module learning path, or click to access them directly from here:

4.3 Assessment Quiz: Basic concepts, representations, and structures of networks
4.4 Small Group Discussion Forum: Revealing peer networks

Reference list

Video 4

This work, "Betweenness centrality", is a derivative of "centrality measures" by Wikimedia user Tapiocozzo, used under CC BY-SA 4.0. "Betweenness centrality" is licensed under CC BY-SA 4.0 by GetSmarter.

This work, "Closeness centrality", is a derivative of "centrality measures" by Wikimedia user Tapiocozzo, used under CC BY-SA 4.0. "Closeness centrality" is licensed under CC BY-SA 4.0 by GetSmarter.

This work, "Degree centrality", is a derivative of "centrality measures" by Wikimedia user Tapiocozzo, used under CC BY-SA 4.0. "Degree centrality" is licensed under CC BY-SA 4.0 by GetSmarter.

This work, "Eigenvector centrality", is a derivative of "centrality measures" by Wikimedia user Tapiocozzo, used under CC BY-SA 4.0. "Eigenvector centrality" is licensed under CC BY-SA 4.0 by GetSmarter.

Video 5

Barabasi, Albert-Laszlo, and Albert, Reka. 1999. “Emergence of scaling in random networks.” Science 286:1-11. Accessed June 22, 2016. doi: 10.1126/science.286.5439.509.

Dong, Xiaowen. 2016. “Networks: Concepts and applications.” Content slides. Massachusetts Institute of Technology (MIT).

This work, "Six degrees of separation", is a derivative of "Experemental Small World (possible option)" by Ageev Andrew, used under CC BY-SA 3.0. "Six degrees of separation" is licensed under CC BY-SA 3.0 by GetSmarter.

Video 6

Dong, Xiaowen. 2016. “Networks: Concepts and applications.” Content slides. Massachusetts Institute of Technology (MIT).

Fortunato, Santo. 2010. “Community detection in graphs.” Physics Reports 486, 75-174. arXiv:0906.0612v2

This work, "Hierarchical clustering", is a derivative of "Hierarchical clustering simple diagram" by Wikimedia user Mhbrugman, used under CC BY-SA 3.0. "Hierarchical clustering" is licensed under CC BY-SA 3.0 by GetSmarter.

This work, "Zachary’s Karate club example", is a derivative of "Karate Cuneyt Akcora" by Wikimedia user Cuneytgurcan, used under CC BY-SA 3.0. "Zachary’s Karate club example" is licensed under CC BY-SA 3.0 by GetSmarter.