Studying Policy-Diffusion: Inference from Text-Edge Networks

Presentation on 13 December 2018 11:00-12:30 h in SR9 at SOWI, Innsbruck. The talk explains how Natural Language Processing and temporal inferential network analysis improve diffusion studies.

AUTHOR

Lisa Lechner

AFFILIATION

University of Innsbruck

PUBLISHED

Oct. 12, 2018

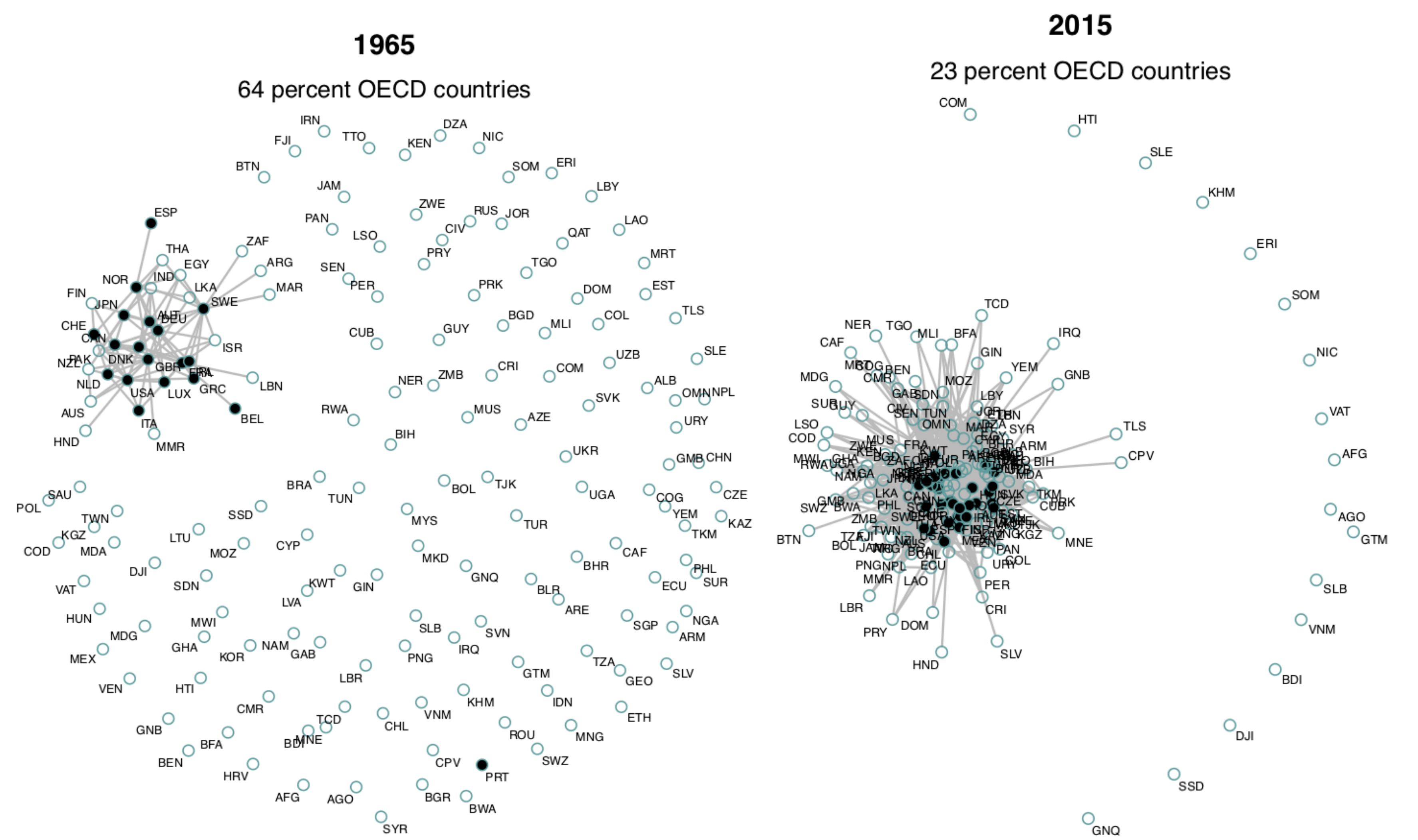

Figure 1: Diffusion of bilateral tax treaty design

> *80 percent of information is text. Authors of texts aim at communicating a message or messages, which are directed to a specific subject, multiple specific subjects, or to a general audience. In short, text connects subjects and creates relational data. Such relationships are often reduced to observation of whether pairs (dyads) are incident to connections. The connection between one dyad is often seen in isolation of the connection between another dyad. Yet, in the real world independence among dyads is unrealistic, because the communication between two actors is likely to depend on communications between other actors. Diffusion scholars study these interdependences. The talk discusses two major challenges in the diffusion literature and how Natural language processing (NLP) and temporal inferential network analysis help overcome these challenges.*

The first challenge for scholars in this area represents the measurement of the diffusion concept. The majority of papers on policy diffusion use a dichotomous measure of policy-(non-)adoption. For two reasons this is problematic. First and as mentioned above, actors communicate the majority of information by dint of text. Squeezing the content of such texts in a yes or no categorization often requires human interpretation and substantial aggregation. Second, such a dichotomous approach increases the likelihood of measuring spurious rather than real diffusion. More fine-grained measures facilitate the identification of policy-diffusion paths. NLP avoids human subjectivity and produces more detailed diffusion measures, which helps taking causal inference more seriously.

The second challenge refers to the model testing for conditions explaining diffusion. Most political scientists analyze the concept of diffusion by dint of regression analysis with spatial lags, which involves high risk of omitted (structural) variables. Dynamic network analysis, which is capable of capturing multiple structural in addition to node and edge terms, is better suited to study diffusion-problems. Moreover, inferential network analysis allows to impose restrictions in the model that make tie formation more realistic. For instance, should scholars account for the fact that actor *j* can only adopt a policy from actor *i* if actor *i* adopted the policy first. Also, one should account for the fact that the wait times between actor *i*'s adoption and future adoptions of other actors are exponentially distributed, which means that short times are more likely than long terms.

In the presentation, I first discuss the challenges in diffusion studies theoretically, then I illustrate these with some examples from my own research, and finally I show the software infrastructure I use to implement NLP and dynamic inferential network analysis in the context of diffusion.