Chapter from the book *New Frontiers in Graph Theory*

Downloaded from: [http://www.intechopen.com/books/new-frontiers-in-graph-theory](http://www.intechopen.com/books/new-frontiers-in-graph-theory)

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Graphs for Ontology, Law and Policy

Pierre Mazzega\textsuperscript{1,2*}, Romain Boulet\textsuperscript{3} and Thérèse Libourel\textsuperscript{3}

\textsuperscript{1}UMR GET, IRD, CNRS, Université de Toulouse III
\textsuperscript{2}International Joint Laboratory OCE, IRD – Universidade de Brasília
\textsuperscript{3}UMR ESPACE-DEV, IRD, Université de Montpellier II

\textsuperscript{1,3}France
\textsuperscript{2}Brazil

1. Introduction

Since the post-war decades, Public Policies are required increasingly as a preferred tool to promote collective action. Today these public policies are developed through sophisticated participatory schemes involving a variety of actors, public or private. Indeed in the current context of globalization though the State and its administration are key actors (Henry, 2004), their influence fades gradually into a more diffuse institutional environment (Östrem, 2005) involving a multitude of other actors (Hassenteufel, 2008). For example at the two ends of the spectrum of governance, we find on one hand the increasing involvement of supranational entities, on the other hand the involvement of nongovernmental organizations.

Therefore, the stated aims of their implementation and the means to achieve them are the result of negotiations and of - at least partial - consensus with stakeholders. In representative governments especially, the trend is growing imposing accountability as regard to the implementation and effects of these policy devices. Following the same trend it becomes mandatory to lean projects of public policy on impact analyses (André et al., 2004; Bourcier et al., 2012) themselves regulated by legal provisions (e.g. European Commission, 2004). These changes induce or reinforce specific phenomena of interest here: a) the systemic nature of law and public policy (and “public action”, an expression that might have a more explicit connotation from a dynamic point of view); b) the gradual - yet limited - disappearance of a singular authority (as could be the State) having the monopolistic power necessary for the shift in or the centralized decision-making; c) the emerging nature of the direct or indirect (desirable or undesirable, internalized or externalized, etc.) effects induced by the policy and law.

These introductory remarks that are part of an intellectual position rather than of a due demonstration, lead us to propose a dialogue between the analysis of public policy and the analysis of complex systems. We take the risk of turning to the analysis of law and public policy as complex systems whose implementation mobilizes law, economics, sociology, policy and administrative sciences, and which understanding relies on these disciplines but

\*Corresponding Author

www.intechopen.com
also, in its foundations, on ontology or mathematics (though we are aware that the scale of
this ambition is balanced by the modesty of the results presented here).

The modeling of law and public policy is an emerging field of research aiming in particular
at understanding their "functioning", at simulating \textit{ex ante} their potential impacts, at finding
ways of achieving a desirable goal (back casting), at producing tools for decision support, at
allowing the participative drafting of legal texts or policy, especially by using ICT
(Information and Communication Technology) or approaches developed for the analysis
and structuring of social networks. Two objectives are achieved in this chapter: a) cover part
of the literature of recent years in this field; b) illustrate our point with original results
drawn mainly from the analysis of law and policy in the wide field of environment and
sustainable development.

In Section 2 we present some basic notions in graph theory and network analysis routinely
used in this article (and in other cited articles). Our choice is to limit ourselves to a very
small number of tools and concepts from graph theory while providing more opportunities
for application in law or political science. In Section 3, an ontology-inspired approach is
used to represent the law and public policy (and the associated knowledge) in the form of
graphs whose analyses are powerful to discover and interpret cognitive and operational
structures in these systems. We hope that the dividend of the example will double as our
analysis focuses on the representation of current research conducted on the application of
ICT in e-governance and modeling of public policies. In Section 4, we present various types
of graphs induced by the law and whose analysis leads to propose a quantitative approach
to some aspects of legal complexity. In particular the analysis of environmental treaties of
international law adds a new graph structure to those previously identified in the analysis
of networks of references between laws.

We proceed similarly in Section 5 for public policies considered in terms of organized
action. It is in this case a policy of management of water resources implemented in France at
the scale of a hydrological basin (large water cycle) where many uses are competing in
particular in period of low water. Some perspective for expanding the areas of graph theory
that will have a great interest in the modeling of Law and policy are briefly mentioned in
Section 6 with concluding remarks.

2. Graphs for the analysis of complex systems

From a general perspective graph theory is today a cornerstone of the study of complex
systems. In this chapter we use a few fundamental concepts, measures and estimation
methods developed in graph theory and network analysis, that provide us with privileged
ways to conceptualize, formalize, model and visualize structural properties of legal and
policy systems and to identify their non trivial properties. Both the use and development of
these concepts and methods have proven to be fruitful in the study of large interaction
networks in sociology or in biological sciences, where emerging structures like small-worlds
or scale-free structures are recurrently exhibited. In return, we hypothesize that the
contemporary legal and policy "fabrics" pose original problems that may trigger new
developments in graph theory and network analysis. In order to show their relevance for the
study of legal and policy systems in the next sections, we restrict our "toolbox" to the
following limited set of concepts and estimators (the analysis methods and algorithms will be indicated through the cited references only).

2.1 Concepts

A graph is a mathematical object consisting of two sets: a set of vertices (or nodes) and a set of edges (or links), an edge linking two vertices. This mathematical structure of graph and its related analysis tools are used to model real phenomena of interaction between objects that are networks (see e.g. Brandes and Erlebach, 2005). We now introduce basic vocabulary of graph theory that will be used in this chapter. An induced sub-graph $H$ of a graph $G$ is a graph whose set of vertices is included in the set of vertices of $G$ and there is an edge between two vertices of $H$ if and only if there is an edge between these two vertices in $G$.

An undirected graph is connected if there is a path between any two vertices. A connected component of a graph is a sub-graph of maximum size (in particular a connected graph has one and only one connected component: itself). A directed graph is weakly connected if the underlying undirected graph is connected. It is strongly connected if there is a directed path (in each direction) from each vertex to any other vertex. A strongly connected component is a maximal strongly connected sub-graph.

A vertex $v$ is a neighbor of a vertex $u$ if there is an edge between $u$ and $v$. The neighborhood of a vertex $u$ is the sub-graph induced by the set of its neighbors and the degree of $u$ is the number of its neighbors. A complete graph is a graph where each pair of vertices is linked by an edge. A clique in a graph is a complete sub-graph. Several matrices can be associated to a graph such as the adjacency matrix $A$ whose $(i,j)$-element is 1 if vertex $i$ is linked to vertex $j$ and is 0 otherwise.

2.2 Measures

Understanding the architecture of a network first passes through a structural analysis. Networks are of many different kinds - simple, directed, weighted, labeled, etc. - which requires the development of specific indices to measure their properties. Some indices measure basic characteristics of the network such as the number of vertices or nodes $n$, the number of edges or links $m$, the density $d$ of the network (defined by the ratio between the actual number of links and the total possible number of links) and the mean degree.

The indices measuring the global connectivity of the network are built on the notion of shortest paths in the network. A shortest path between two vertices is a path (a sequence of consecutive edges) of minimal length. Global connectivity is then evaluated by the mean of the shortest paths lengths $l$, the characteristic path length $L$ (median of the means of shortest paths; Watts, 2003) or the diameter $D$ (the length of the longest shortest path in the graph). Two indices, called clustering coefficients, measure the local connectivity. The first clustering coefficient $C1$ is defined by the average density of the neighborhood of a vertex. The second clustering coefficient $C2$ is the ratio between the number of triangles in the network and the number of connected triples in this network which can also be seen as the probability to have a link between two vertices linked to a common vertex.
In order to evaluate the position of a vertex in a graph, three notions of centrality have been defined (Freeman, 1979). The degree centrality $C_D$ measured by the degree of a vertex, the betweenness centrality $C_B$ related to the number of shortest paths going through a vertex and the closeness centrality $C_C$ of a vertex based on the mean distance to the other vertices. Then the centralization of a network is the normalized sum of the differences between the vertex with the highest centrality and the centralities of the other vertices in the graph. A zero (resp. unit) value of the network centralization corresponds to a non centralized (resp. most centralized) network. The degree distribution is the probability distribution of the vertex degrees; for each integer $k$ it gives the probability to have a vertex of degree $k$. Other indices may be introduced dealing more particularly with the impact on the network structure of weights or the presence of different kinds of links (Boulet, 2011; Boulet et al., 2011b).

2.3 Finding communities

An essential step in network analysis is the research of communities. In a social network, some people can be gathered into communities. Following this example of community we can extend this notion to networks of different nature (networks of legal texts, lexical networks, etc.) analogously defined as groups of vertices more densely connected to each other than the rest of the graph. Detecting communities in networks is a very active field of research in network analysis and we shall consider only a few of the existing algorithms.

Several algorithms have been developed relying on various methods. The walktrap algorithm (Pons and Latapy, 2005) is based on random walks on graphs with the underlying idea that a random walk would be trapped into a community. The fast greedy algorithm (Clauset et al., 2004) aims at greedily maximize the modularity (a criterion assessing the quality of a partitioning) of the resulting partitioning. Finally spectral methods (von Luxburg, 2007) are based on eigenvalue decomposition of matrices associated to a graph which embeds the graph in an Euclidean space on which we can use statistical clustering.

A special community is worth mentioning here, the rich-club (Zhou and Mondragon, 2004). A rich-club is formed when the vertices with highest degree (the rich vertices) are highly interconnected and form a very dense group (a club). They therefore constitute a central and influent community in the network.

2.4 Types of graphs

An Erdős-Rényi random graph (Erdős and Rényi, 1959) is a graph where an edge is put between two vertices with a uniform probability $p$. It is known (Watts, 2003) that such a graph have a tight global connectivity (the randomness of edges creates shortcuts) and low clustering coefficients (the local density looks like the overall density). Then the indices presented earlier allow distinguishing several classes of graphs. A small world network (Watts, 2003) is a network with a tight global connectivity (nearer than that of a random graph) and a high local connectivity (much larger than that of a random graph). A concentrated world (Boulet et al., 2011a) is a dense graph with some highly interconnected vertices having the highest centrality measures. A scale-free network is a graph with a power-law degree distribution. Therefore it has a lot of vertices with a low degree and few vertices with a high degree (sometimes called hubs). It contrasts with Erdős-Rényi random graphs, the degree distribution of which is centered to a mean value and follows a Poisson
distribution. A directed graph is a graph with directed edges. A special kind of directed graph is a tournament in which there exists an edge between any two vertices (it can be seen as a complete directed graph).

3. Ontology as a network

The title of this section calls for comments. First, even if we exclude *a priori* from the scope of our discussion the field of philosophy, the term "ontology" has no unambiguous acceptance. By the 1980s, the field of computer engineering and knowledge representation takes the term to define a computational model for performing automated reasoning. In 1995, T. Gruber says that "an ontology is a formal, explicit specification of a shared conceptualization" (Grüber, 1995). In 2009, he refines his remarks by stating that "an ontology defines (specifies) the concepts, relationships, and other distinctions that are relevant for modeling a domain. The specification takes the form of the definitions of representational vocabulary (classes, relations, and so forth), which provide meanings for the vocabulary and formal constraints on its coherent use" (Grüber, 2009). Finally, the pragmatic issues converge in the statement that “an ontology is a tool and product of engineering and thereby defined by use ». The key issues opened by these proposals relate to how to make the best choice of language, and how to set objectives of use of these ontologies.

3.1 Languages and objectives for ontology

In terms of languages, it is interesting to quickly scan the history. Originally, we find the semantic networks (Quillian, 1968) which are doubly labeled directed graphs, the vertices being labeled by concepts, the arcs being labeled according to relators between concepts. Then follow, the most significant proposals emanating from Minsky, Sowa, Brachman and Levesque. Minsky (1975) proposes an approach based on the notion of a grain of knowledge or frame and specifies that « here is the essence of the theory: when one encounters a new situation (or makes a substantial change in one’s view of the present problem) one selects from memory a structure called a frame. A frame is a data-structure for representing a stereotyped situation. Attached to each frame are several kinds of information. We can think of a frame as a network of nodes and relations. The "top levels" of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many terminals—"slots" that must be filled by specific instances or data».

Sowa (1984) proposes the conceptual graph formalism. Those are bipartite graphs with two types of nodes: concept-nodes and relation-nodes. Knowledge representation can be defined in a specific area, this being reflected in what is referred to as a support and includes two lattices: the concepts and relationships pertaining to the chosen field. The reasoning is based on semantics in first order logic, but a more original view suggests compensating the mechanisms of logical inference by graph operations in particular the homomorphism of graphs (called projection; Chein and Mugnier, 2009). Brachman (1979), Brachman and Levesque (1985, 2004) propose extensions to the language of frames and semantic networks from semantic-based description logics. They introduce the notions of concept, role and individual and rely on an inference mechanism based in particular on subsumption.
These theoretical propositions are the heart of various studies conducted to date. They are accompanied by proposals for operational developments, the most iconic of them being probably KIF (Knowledge Interchange Format) on the initiative of the consortium DARPA Knowledge Sharing Effort (Genesereth, 1992), the Ontolingua project (Gruber, 1992) until the arrival of the Semantic Web following the original article by Berners-Lee et al. (2001), which paves the way for many of the proposals in terms of protocols, languages, standards (XML Extensible Markup Language, Resource Description Framework RDF, RDF Schema, OWL Web Ontology Language) and tools. In the field of legal studies these researches are now flourishing (Casanovas et al., 2007; Sartor et al., 2011).

Let us return to the question of the use and purpose of ontology design in the context of knowledge of public policies. It is now commonly accepted (Guarino, 1998; Gangemi et al., 2001) that ontologies can be stratified by level of generality or abstraction (Figure 1): a) top-level ontologies describe very general concepts like space, time, material, objects, events, actions, etc., which are independent of one problem or particular application domain. They specify very general abstract knowledge whose content depends on the degree of formalization; b) domain ontologies and problem-solving ontologies or task ontologies that describe, respectively, the knowledge related to a generic domain (like medicine or aviation) or generic activities (as water management) by a specialization of the concepts presented in the top-level ontologies; c) application ontologies describe concepts depending both on a field and on a particular task. They are specializations of the two previous types of ontologies.

![Fig. 1. Scheme of the stratification of ontologies by levels of generality and abstraction.](image)

Overall, it appears that the graphs are essential in the development of knowledge representations and ontologies, and that the synergy between different approaches is important (Farquhar et al., 1995). The example we now develop tackles with the representation of a field of knowledge. The approach taken may not be consistent with the *doxa* but illustrates the types of exploration that allows the use of graphs in the field.

### 3.2 Classification and phenomenological graphs

Let us return to some fundaments. We here call “classification graph” a graph built on the relationship of subsumption ("is a" relationship) that binds an instance to its class...
membership. According to an Aristotelian-type classificatory approach, a concept (e.g. an agent of the administration) may be both a special case - one instance - to a larger class (the agents) and a (sub-) class for other more specific instances (e.g. the prefect). The support structure of these univocal classifications is of course the tree-graph that we propose to call "class graph". According to its purpose, this structure is effective and unambiguous.

But by destroying any flexibility of interpretation, this normativeness may be undesirable. Let us consider an example which is specific to Law. The judge is asked to reinterpret terms or concepts in a manner more consistent with the acceptance that tends to prevail in a constantly evolving social context. Thus the "family" of the Civil Code of 1804 is no more precisely that family of a society that tends to recognize the single-parent family, same-sex parents or surrogate mothers. Sometimes it follows a reassessment of existing case law, the judge taking a unique position that other judges will adopt later. For schematic that can be our example (indeed, the departure from precedent is less about the interpretation of terms than about the reinterpretation of a point of law which, presumably, involves an entire ontological environment), a sociologically-oriented ontology should be able to allow, in the organization of its classes and of their instances, the occurrence of multiple relations of subsumption (one instance being subsumed to several classes), in a way more conform to uses than in line with normativeness of any a priori model (should it be dominant).

Incidentally, we move from the trees (and forests) to the wider world of all graphs. These graphs then capture the structure of a conceptualization of a system that relies on a variety of types of objects and typed relations, a "model" or "ontology". In this chapter we propose to call these structures "phenomenological graphs" - or "pheno-graph" for short - because they capture the observable structure (whatever the means of observation) of relationships between entities whose existences are attested.

There are various ways to build such a model, based on pragmatic approaches in the fields of law or public policy. One of them is of particular interest in this section. It is based on the terminological analysis of large corpus of texts (legal texts, texts describing public policies, regulations, etc.). The analysis of terminology and text mining can also be based on a representation of the knowledge domain in the form of ontology, the text analysis revealing more or less reinforced relationships between terms or concepts (or even invalidating the relevance of some ontological relations in this context or by inducing new links).

Formally, this type of analysis consists of two levels say a tree-like classification for supervision and the texts revealing a more general network of entities linked by heterogeneous relationships (such double structure is described in Mazzega et al., 2011). Let us now work an example of a double structure - ontological and phenomenological - in the field of research mapping.

3.3 Illustration: Research in e-governance & policy modeling

The CROSSROAD European project produced a public report on the state of the art on the field of ICT research in electronic governance and policy modeling (CROSSROAD, 2010) from which we here extract and analyze the data structure. This research field is divided in five domains divided in areas and sub-areas as follows: 1) Open government information
and intelligence for transparency (4 areas; 18 sub-areas; associated color: orange); 2) Social Computing, citizen engagement and inclusion (4 areas; 14 sub-areas; green); 3) Policy modeling (4 areas; 18 sub-areas; yellow); 4) Identity management and trust in governance (4 areas; 18 sub-areas; blue); 5) Future internet for collaborative governance (5 areas; 25 sub-areas; purple). In this list we attribute a color to each research domain and to its afferent areas and sub-areas for the purpose of graph visualization.

Fig. 2. A representation of the pheno-ICT graph with the size of each vertex being linearly related to its betweenness degree (see text). Vertices of the same color belong to the same research domain.

The full mapping of this field of research shows the type of double structure mentioned above, say: a) a tree structure - that we call onto-ICT graph (not represented here) - induced by the organization of this field of research in the above mentioned 5 domains (see below), 21 areas and 93 sub-areas (grouped under the term "research topics") that are the 119 vertices of the graph, with 114 edges; b) a network structure – which principal connected component we call pheno-ICT graph (111 vertices, 245 edges) - in which topics (research domains, areas or sub-areas) are linked by pairs when the specific researches they cover are explicitly related in a noticeable amount of scientific publications or when these publications refer to specific terminologies preferentially associated with these topics.
In Figure 2 we give a representation of the *pheno-ICT graph* that enhances the relative magnitudes of betweenness centrality associated to the vertices. The correspondence between the labels of the vertex (“x” type label for domain, “x.x” for area, “x.x.x” for sub-area) and research field they cover can be found in CROSSROAD (2010). Let us now imagine some simple scenario (about research policy for policy research). Should we organize some meeting in ICT research in electronic governance and policy modeling, what should be the title of the sessions? A convenient choice might be to choose those topics with the highest degrees in the *pheno-ICT graph*. If now some (financial or human) resources are available to strengthen research where should it be primarily allocated? Such a choice of course depend on the criteria for decision making but a strategy could be to support those topics that present a high betweenness centrality, because they will reinforce the synergy between various research communities and strengthen the whole research system.

<table>
<thead>
<tr>
<th>Label</th>
<th>Name of the research topic</th>
<th>Betweenness Centrality Index (rank)</th>
<th>Normalized Degree (rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Social Computing</td>
<td>1.00 (01)</td>
<td>0.64 (05)</td>
</tr>
<tr>
<td>3.1</td>
<td>Policy Analysis (in policy modeling)</td>
<td>0.86 (02)</td>
<td>1.00 (01)</td>
</tr>
<tr>
<td>3.3</td>
<td>Visualization (in policy modeling)</td>
<td>0.80 (03)</td>
<td>0.75 (03)</td>
</tr>
<tr>
<td>2.2</td>
<td>Citizen Engagement</td>
<td>0.78 (04)</td>
<td>0.46 (06)</td>
</tr>
<tr>
<td>2.3</td>
<td>Public Opinion Mining &amp; Sentiment Analysis</td>
<td>0.60 (05)</td>
<td>0.39 (07)</td>
</tr>
<tr>
<td>1.3</td>
<td>Visual Analytics (in open government information)</td>
<td>0.36 (06)</td>
<td>0.39 (07)</td>
</tr>
<tr>
<td>1.2</td>
<td>Linked Data (in open government information)</td>
<td>0.26 (07)</td>
<td>0.14 (13)</td>
</tr>
<tr>
<td>4.3</td>
<td>Trust (in governance)</td>
<td>0.16 (08)</td>
<td>0.11 (17)</td>
</tr>
<tr>
<td>1.3.3</td>
<td>Analytical Reasoning (in visual analytics)</td>
<td>0.14 (09)</td>
<td>0.07 (22)</td>
</tr>
<tr>
<td>4.1</td>
<td>Identity Management</td>
<td>0.12 (10)</td>
<td>0.36 (09)</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Web Accessibility (in human / computer interactions)</td>
<td>0.11 (11)</td>
<td>0.07 (22)</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Public Service Aggregation, Mash-ups &amp; Orchestration</td>
<td>0.10 (12)</td>
<td>(&gt;25)</td>
</tr>
</tbody>
</table>

Table 1. Ranking the first twelve research domains, areas or sub-areas of ICT research in electronic governance and policy modeling by decreasing (normalized) betweenness centrality index (as extracted from the pheno-ICT graph) considering only entering links. The normalized degree (and rank) of the research topics are also given for comparison.

In Table 1 we give the first twelve research topics ranked by decreasing (normalized) betweenness centrality index. The normalized degree (and corresponding rank) of the research topics are also given for comparison. It appears that though they do not have a high rank of connectivity, several topics have a high rank in terms of betweenness centrality. This is the case for the first topic “social computing” (betweenness centrality rank=1; connectivity rank=5) that is about the application of web social software in public-sector activities and the support to social interaction and collaboration (and with the associates terminology: social networking, content syndication in government portals, blogging and
micro-blogging, collaborative writing tools, feedback, rating and reputation systems –
CORSSROAD, 2010). It is interesting to notice also that several topics among the most
connected with the other topics are not in the short list of high betweenness centralities: area
1.1 “open & transparent information management” with degree 0.89 (2\textsuperscript{nd} rank), area 3.2
“(policy) modeling and simulation “ (degree 0.75 – 3\textsuperscript{rd} rank), area 5.3 “multi-channel access
and delivery of next generation of public services “ (degree 0.32 – 10\textsuperscript{th} rank), areas 1.4
“findings (in open government information and intelligence for transparency)” and 5.1
“cloud computers (for collaborative governance)” (degree 0.21, 11\textsuperscript{th} rank).

<table>
<thead>
<tr>
<th>General Characteristics</th>
<th>Index</th>
<th>Non-oriented pheno-ICT Graph</th>
<th>Average Indices of Erdös-Rényi Graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>111</td>
<td>109.56</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>239</td>
<td>239.08</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>0.039</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>4.31</td>
<td>4.36</td>
</tr>
<tr>
<td>Global Connectivity</td>
<td>l</td>
<td>3.25</td>
<td>3.34</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>3.15</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>8</td>
<td>6.92</td>
</tr>
<tr>
<td>Local Connectivity</td>
<td>C\textsubscript{1}</td>
<td>0.34</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>C\textsubscript{2}</td>
<td>0.12</td>
<td>0.038</td>
</tr>
<tr>
<td>Network Centralization</td>
<td>C\textsubscript{D}</td>
<td>0.24</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>C\textsubscript{B}</td>
<td>0.11</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>C\textsubscript{C}</td>
<td>0.27</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Table 2. Small world indices calculated on the non-oriented pheno-ICT graph and for
comparison on a set of 1000 Erdös-Rényi random graphs with the same size and order. n is
the number of vertices, m the number of edges, d the density, k the average degree, l the
length of the average shortest path, L the characteristic length, D the diameter, C\textsubscript{1} the first
clustering coefficient, C\textsubscript{2} the second clustering coefficient, C\textsubscript{D} the degree centrality, C\textsubscript{B} the
betweenness centrality and C\textsubscript{C} the closeness centrality (see Sec.2).

On one hand these simple findings can be useful for designing a research policy in this field,
as briefly suggested by our quite trivial scenarios or by many others related to real life
situations. On the other hand it is interesting to see if the pheno-ICT network belongs to a
known class of graphs. We consider the undirected graph (some geodesics do not exist on
the directed graph) for which we assess the general indices as well as the global and local
connectivity (Table 2). The random model used for comparison is the G(n,p) model of
Erdös-Rényi (See Sec.2) where there is an edge between two vertices with a probability p equal to
the density of the graph. The examination of these connectivity indices leads us to conclude
that this network is a small world. Indeed, it fills the following two conditions: a) tighter
global connectivity similar to that of random graphs; b) strong local connectivity, much
higher than that of a random graph. So we are dealing with a type of graph which shows the
structure of many social networks the study of which is enriched with other concepts
opened to interpretation, but that we will not discuss in this chapter.

We also observe that the (non-oriented) pheno-ICT network exhibits high values of the
betweenness centrality (but not of the closeness centrality) when compared to random
graphs. We interpret this indicator as showing that the work developed in many topics
(domains, areas or sub-areas) are routinely called upon to involve or interest not only directly related topics but also other neighboring topics albeit slightly more distant (on the network). In a sense, this feature reflects the strong identity and overall coherence of research in the field of ICT research in electronic governance and policy modeling.

4. Law-induced networks

According to the classic positivist theory of law (knowing that we do not consider here the Common Law), legal norms must obey a hierarchical organization implying that the norms of lower degrees should not be in conflict with the norms of higher degree (Kelsen, 1960). At the top of the pyramid of norms is the Constitution, more or less expanded with a “block of constitutionality”, followed by the organic and other laws, decrees, etc. With the rise of Community, European and International law, also with the proliferation of sources of law, this hierarchical organization is deeply upside. The doctrine echoes these on-going transformations (eg. Ost & van de Kerchove, 2002; Delmas-Marty, 2007) that are also subject of much debate in the democratic representative bodies.

For example in France, the positions of the Court of Cassation and of the Council of State have been for nearly 15 years in opposition as to the primacy of European law over national law, until the adoption of a common position recognizing it during a reversal of precedent of the State Council (see Bécane et al., 2010). Of course this rule is accompanied by measures not to amputate the representative bodies of their ability to legislate or to exercise control over norms produced outside their assemblies, the national legislature thus acquiring a role in the European legislative process (Article 88-4 of the Maastricht Treaty).

So we perceive that under the combined effects of globalization and the empowerment of communities and organizations (e.g. NGOs) network structures appear in all domains of Law and policy. Different kinds of networks: institutional networks (the link being instantiated by roles, powers, information flows, etc.), networks associated to the process of decision-making (e.g. EU), networks and flows in law-making, etc. But outside of these role and power games the analysis of which is subtle, another phenomenon reflecting the overlapping of legal systems can be easily observed and studied. Although its range is relatively minor (it is of interest for some operational services of law broadcasting like LEGIFRANCE www.legifrance.gouv.fr/, EURLEX eur-lex.europa.eu/en/index.htm, etc. and also for codification), its analysis allows starting the development of measures of certain manifestations of "legal complexity": this phenomenon is the emergence of networks of citations between legal texts (Bourcier and Mazzega, 2007a; Bommarito and Katz, 2010). The citations between legal texts results from a self-organizing process, no top authority being in charge to manage the constant flow of legislative regeneration, neither at the national nor regional or global level.

Based on previous work, we will support the next two assumptions that do not deny the concerns of the science involved in the analysis of complex systems, namely: a) that the types of structures found in the network of citations depend on both the size of the object in question and the scale of resolution for their analysis; b) that the "canonical" representation of a field of knowledge - in this case the legal knowledge - deserves to be coupled with a representation based on the analysis of emerging phenomenological graphs. We will then give further arguments in favor of these two hypotheses, based on an empirical analysis of the bipartite network of environmental treaties of international law.
4.1 Scale matters in law

Let us start with small scales. The article is classically considered as the smallest entity that deployed in a broader frame – law, code - should have a legal content making sense of its own. Yet we have found up to 4 levels of subdivision in a significant number of articles in the environmental code, and we have estimated their frequency distribution and length of text statistics (Bourcier and Mazzega, 2007b). The association of a graph of citations to a corpus relies on the choice of the resolution for analysis: should we keep all the subdivisions or stop at the granularity of articles (with 1266 of them in the environmental code, legislative part)? This choice is in fact dictated by the objectives of the thematic analysis. However we understand intuitively that the absorption of certain entities (e.g., subdivisions of articles) in the entity of which they are parties (e.g. article) increases the density of the network by aggregating the parts in the local whole and assigning all edges to it.

On the other side of the spectrum of sizes, the choice of a definition of the analyzed object is just as important. The environmental code has in a sense a unity conquered by the legislature (especially with regard to the rural code) but partly arbitrary since it cites (or is quoted by) 35 other codes, several EU Directives, international treaties etc. Considering at a given time all of the existing law in the world and its ramifications (to which corresponds a global legal-graph), being interested only in the French environmental code, or in all of the codified French law (which associated graph we call hyper-code), is equivalent to extract an induced sub-graph and analyze it separately. This sub-graph can be seen as associated to an ontology-derived community which does not have the emerging character of the phenomenology-derived communities we will consider below. Indeed the ontology-derived communities are deliberately constructed by the legislature, or in this case, coder (two “actors” that deserve being considered themselves as communities of actors).

The use of approaches presented in Section 3 also allows identifying vertices with remarkable properties: for example, articles (resp. Codes) of high degree centrality or high betweenness centrality in the environmental code (resp. hyper-code graph; Mazzega et al., 2009; Boulet et al., 2011a). Thus, as in many complex systems the choice of the analysis resolution and of the size of the object changes the characteristics and properties of the associated graphs, but also the graph type and hence the paradigm of interpretation.

4.2 Alternative representation of legal knowledge

The matter of a code - as the Environmental Code - is usually organized into divisions - books, titles, chapters, sections, articles - that form a tree structure (table of content). However, another organization is behind it, which brings together the divisions with a high density of inter-citations (the main connected graph component associated with the legislative part of the Environmental Code has 980 vertices and 2186 edges). The use of multiple algorithms for partitioning the graph associated to a corpus, allows identifying stable (say found by all algorithms based on different criteria for partitioning - cf. Sec. 2.2) "hidden" communities of its divisions - including articles - which are semantic units pertinent for the interpretation of the law.

Two points are worth noting: a) the existence of these phenomenological communities is known neither to the legislature nor to the lawyer, b) some of these emerging communities
do not coincide with the ontological divisions, necessary in the planned organization of the code (Boulet et al., 2010). Conversely some ontological divisions are not in any stable community, such as for example the provisions relative to New Caledonia or Antarctic (a result that we keep to interpret). In other words, the codified substance could be distributed in a different cognitive representation and determined according to objective criteria (reproducible between codes - possibly from different countries).

Similar results are obtained when analyzing the hyper-code graph of citations among all codes of the French legal system produced using the same codification methodology and whose vertices are the codes (52 to date). Thus we have mainly highlighted (Boulet et al., 2011a,b): a) a rich-club (see Sec. 2.3; Colizza et al., 2006) including the 10 most central and influential codes of the French legal system; b) a stable community of 12 codes governing the legal areas linked to social activities, their regulation and security; c) a second stable community of 11 codes governing the legal areas related to land management, the territories and their resources. Only the analysis of hyper-code code could lead to identify those communities whose existence should challenge the doctrine, their existence being previously unknown.

4.3 Illustration: A tournament of international treaties

We now consider the state (as of 2010) of ratification of the 42 environmental international treaties by the 196 countries (or entities of equal status, such as the European Community) members of the United Nations (see the Chapter XXVII of the database of international treaties of the United Nations http://treaties.un.org/Pages/ParticipationStatus.aspx). Initially we built a graph with a link between countries A and B if both have ratified at least one common treaty. The links are then weighted by the number of common treaties ratified by the two countries. The partitioning of this graph showed two groups (Boulet et al., 2012): one group of 38 countries with only Canada being not part of continental Europe, and a group bringing together 158 countries.

In a second step, we seek a graph showing a higher discriminatory power of any developed strategies for ratification by member countries. For this we used the following rule: a link is established from country B to country A if B is the first country to ratify a particular treaty after ratification of that treaty by A. A second rule must be established with respect to N countries simultaneously ratifying the same treaty: a) link the N countries, which induces a clique in the graph or b) not link the N countries, which induces a stable graph. In the following analysis we use option (a) as our second rule. The graph (196 vertices, 2832 links) we obtain is shown in Figure 3. The hypothesis that motivates the analysis of this graph is that the sequence of dates of ratification reveals (whether collaborative or competitive) collective political strategies of countries (Louka, 2006).

The five countries most ready to sign the environmental treaties are (ranked by decreasing order of the degree centrality) Norway, Finland, Hungary and at the same level, France and Luxembourg. The countries with the highest betweenness centrality, and thus being somewhat in the stream of the temporal succession of ratifications are (ranked by decreasing order of betweenness centrality) Hungary, France, Belgium, Romania and Norway. These simple results show unambiguously the active role of Hungary, Norway and France about the environment on the international stage.
We find four stable communities with 41, 39, 22 and 10 countries respectively. Most European countries (in a geographical sense) are the second community, alongside the United States, Canada, Japan, New Zealand and Burkina Faso. The contrasting positions of these countries with regard to Europe, in particular on issues related to the climate change and energy policy suggests that this grouping into a single community rather reflects an antagonist-type of interaction. Note that this community also has the largest clique of the graph composed of 18 countries (Germany, Austria, Belgium, European Community, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, Portugal, United Kingdom of Great Britain and Northern Ireland, Slovakia, Sweden).

Among the emerging countries, Brazil, China and India are in the third community (22 countries), but not South Africa (in the 4th community). The Russian Federation does not appear in any stable community (as defined here in Sec. 2.3), as also 83 other countries.

Fig. 3. A representation of the oriented graph associated with the ratifications of the environmental treaties and conventions of International Law. The vertices are 196 countries (or analog legal entities like the European Community) of the world. A link is drawn from B to A if B is the first country having ratified a given treaty just after country A. The larger orange (resp. smaller green) vertices represent countries with the highest (resp. lowest) betweenness centrality (see text).
Countries suffering from acute chronic internal conflicts, backed to failing States, are among the countries that have ratified the least environmental treaties, a fact that reveals a simple inspection of their degree centrality. The logic behind the arrangement of the first community, made up of many countries from different geographical, geopolitical or economic regions, is not easy to identify. This probably results from the existence of several smaller sub-communities but whose ratifications are interspersed.

So if these early results are promising (as we think about the other 28 chapters that bring together hundreds of treaties), a more detailed analysis of these data, taking into account the particular dates of ratification themselves and a diachronic knowledge of the major world events, should identify more subtle patterns of international relation strategies as expressed by the evolution of international law (Schneider & Urpelainen, 2012).

5. Graphs for policy analysis

The analysis of citation networks between legal texts foreshadows network analysis for components and links of heterogeneous networks that may be associated with public policy. We have already mentioned that throughout the life cycle of policy (agenda-setting, formulation, decision making, implementation, evaluation; Howlett and Ramesh, 2003) are mobilized a variety of actors (individual or group, institution, etc.) and of resources.

One objective of these models is to provide tools for the ex-ante assessment of the potential impacts of these policies. This ambition is very high and we believe should be seen as a horizon to strive for rather than an achievable goal in the near future. Indeed at the individual scale, actors' behavior is unpredictable, their resilience and creativity never limited even if in a highly regulated or binding frame. However, at the collective level, patterns of behavior are observed and can be analyzed using proven methodologies. The analysis of graphs associated with the representation of certain aspects of public policy adds to the available (or developing) tools promoting a better understanding of these instruments of the "public" collective action. Rather than remain at this level of generality we prefer in this section to develop an example who keeps illustrative in general. This is the development of a simulation platform of the impact of new norms for managing water resources at the scale of a hydrological basin.

5.1 A framework for water policy

The water management in France is governed by various legal provisions, including the Law on Water and Aquatic Environments 2006 (LEMA, 2006) which adopts tools to achieve by 2015 the goal of "good" water state as set by the European Water Framework Directive (WFD, 2000). Basically the water management is organized at the scale of large hydrological basins. Southwest of the French metropolitan territory, the Adour-Garonne basin (AGB) covers about 120 000 km², with a rate of precipitation of 600 mm to 2000 mm/year, a potential for runoff of 90 km³/year and large stocks of water in the aquifers. About 7 million people are distributed over 6900 municipalities and 35 cities with more than twenty thousand inhabitants (most such data are provided by the website of the Adour-Garonne Water Agency www.eau-adour-garonne.fr/).
Average annual withdrawals are 2.5 km³, being roughly equally distributed between the three main uses, namely drinking water, industrial uses and agriculture. However, the water availability and uses are very uneven over the year: during the low water period (summer and early autumn) agricultural levies represent 80% of the withdrawals for about 645 000 irrigated hectares. During the heat wave of 2003 the withdrawal for irrigation has nearly doubled compared to average years. On the other hand, the climate change scenarios predict that 2003 will be in terms of rainfall better than the average year of the 2050’s (Pagé & Terray, 2011). Consequently, the management of low water is the first priority of the Adour-Garonne Water Agency. The water uses must also be balanced with the preservation of aquatic ecosystems and environmental services, two issues that the EU directive and the LEMA (2006) explicitly raised as priorities.

On the scale of the Adour-Garonne basin, the water policy and its implementation was prepared with the participation of governmental administrations, territorial and other communities, associations, the civil society, etc., over several years, as described in the master water planning and management report (SDAGE AG, 2010) and in the associated implementation program (PDM AG, 2010), documents whose provisions may be legally opposed to other sectoral policies (territorial development, economic development, etc.). More locally, across the river basin management units (sub-basins), the low-water management plans specify the measures to regulate water uses within the environmental, social, economic and ecological constraints associated to periods of low flow.

In the MAELIA project (Multi-Agents for EnvironmentaL norms Impact Assessment) we develop a hybrid model combining a multi-agent system, model equations and GIS to estimate the societal, economic and environmental impacts of new “norms” implementation in the basin. We focus specifically on the new device that will limit water withdrawals for irrigation in many watersheds. To evaluate ex ante the potential impacts of this new frame, we represent in the model both the behavior of key actors involved in the water uses or management, but also the environmental (including hydrology) and climate dynamics as well as the main activities (especially agriculture) developed by the actors that may have a direct or indirect effect on the resource. To achieve this goal it is first necessary to identify the key actors, the activities they lead, the resources (material and cognitive) they use or generate, and the social and environmental processes the entire system depends on (Sibertin-Blanc et al., 2011; Thérond et al., 2011). This step was carried out by an interdisciplinary participatory method that results in the production of a conceptual model presenting only the essential elements of the management of low water in the basin.

### 5.2 Illustration: Basin-scale low water management

The conceptual model is built up on the basis of various UML diagrams (OMG, 2005). Figure 4 shows the main one, namely the actor- resource diagram (see Sibertin-Blanc et al., 2011 for some other diagrams). The vertices are instantiations of the following classes: actors, material resources, cognitive resources (including the norms: legal texts, programs, directives, etc.). The links are of different types (with cardinalities that we discard here): actors’ activities, social processes, environmental processes or conceptual relationships between the linked entities. We are interested in the structure of the graph associated to this conceptual model of the low water management in the Adour-Garonne basin.
The graph has a single component of 66 vertices and 87 edges. It is weakly connected, with only one strongly connected group of vertices formed of two material resources (equipment for drinkable water catchment; equipment for raw water catchment and flow like canals, pipes, etc.), two “actors” (remembering that the term « actor » can refer to a group of actors: raw water distributors; clean water distributors) and one cognitive resource (structure of water prices and charges). Functions or operations associated with the vertices with the highest degree of connectivity must be modeled accurately because they directly affect the states of many others. These are in particular (by decreasing degree) the farmers, the (public or private) equipments, socio-professional categories (regarded as consumers of drinking water in municipalities), all types of water tanks, the land sub-parcels.

Fig. 4. The graph associated to the conceptual model of the law water management in the Adour-Garonne basin (France). The colors of the vertices indicate their type: actors (orange round rectangles); cognitive resources and norms (green circles); material resources (yellow triangles). The labels, attributes and functions associated to each vertex in the UML model have been removed to clarify the graph (some of them being given in the text).

On the distribution drawn on Figure 5 we see that around ten vertices present a significant betweenness centrality index. These vertices are the following ones: (label=49) Raw water distributors; (09) Manager of engineering structure (dams, etc.); (61) Manager of water catchment; (47) Clean water distributors; (57) Water tanks; (07) (Public or private) Equipments; (14) Equipment for water catchment; (17) Equipment for drinkable water catchment; (00) Geo-referenced territorial unit; (01) Plots. From a perspective of how the dynamical model works, the behavior of these entities (including changes in their respective states and comparing them with empirical data) should be given special attention. Indeed their intermediate position a priori designate them as being likely to exhibit abnormal
behavior in the case all processes and activities modeled are not adequately represented or adjusted against each other.

A look at the global and local connectivity indices estimated for this graph does not allow us to characterize it as a small-world network. Although the overall connectivity is tightened (L = 4.14, D = 8) and close to that of random graph (L = 4.58, D = 11.32), the local connectivity is low (C1 = 0.41, C2 = 0.027): the first clustering coefficient is artificially high due to the large number of vertices of degree 1, the second clustering coefficient is low. Moreover, the graph contains only two triangles.

![Fig. 5. Distribution of vertices of the low water management graph (Fig.3) in descending order of their degree of betweenness centrality.](image)

All these results concern the static structure of the conceptual model of low water management. From the conceptual model is then derived a dynamic model where all the social and environmental processes (represented by computer codes) change over time the state variables (or attribute values) of the actors and resources. The analysis of the dynamic properties of this type of model must use a range of other tools (Mesbahi and Egerstedt, 2010), including statistics we will not address here, but whose scope covers any multi-agent system, being dedicated to the modeling of public policies and of their potential impacts or to any other complex system dynamics.

6. Conclusion

As was briefly shown in this chapter network analysis is useful for many purposes in Law and Policy: to identify emerging cognitive sets or patterns (groups of concepts); to make a functional mapping of the components of a complex system and of their interaction; to carry out all stages of modeling (from the definition of a referential representation to the setting up of a conceptual model and its implementation in a simulation platform for impact assessment); to analyze social dynamics (like strategies, coalition dynamics, etc.).
Many other concepts can be requested from graph theory (Jungnickel, 1999) to highlight the hidden properties of legal corpuses or public policies. For example we mentioned in Sec.4 the impact of the aggregation of nodes (corresponding to different choices of resolution) in a root vertex on the density and other measures associated to a network. A similar transformation occurs when passing from a given graph to the graph of its stable communities. Graph theory can provide fundamental knowledge on the incidence of these operations on the main measures associated with the studied networks, which would constitute a major theoretical and methodological contribution to the analysis of complex legal and policy systems.

As soon as graphs are involved in formal representations, their comparison provides a methodology for comparing the complex legal and policy systems they capture. Such issues appear especially in ontology alignment, in the identification and analysis of functional similarity of actors’ positions in governance schemes, in the comparison of policies or legal systems between different sectors or countries, etc. Techniques of graph mining (Cox and Holder, 2007) would be particular useful in conjunction to text mining, a huge amount of data being accessible today via institutional public web sites and data base systems (as illustrated in this chapter with LEGIFRANCE, EURLEX or the UNO data basis on the multilateral treaties).

Another exciting area of research is to provide a graph-related equivalent to concepts developed and tested in the social sciences, especially law and political sciences. Different schools of thought have developed conceptualizations of central concepts such as power or role (among many others), which could be translated in terms of graphs and their properties and thus be enriched by the theoretical contributions of mathematical research.

Finally, the legal and policy “factories” are in constant work. The objects they produce continuously change, reform themselves, aggregate or distinguish each others. The approach to these phenomena by graphs that are either the support for dynamics (Egestedt and Mesbahi, 2010), flows or events, or even whose topology changes is therefore essential.

We can even imagine that these formal approaches based on graph theory will soon be integrated in the design of inter-sectoral integrative policies, thus responding to a growing need for multi-level societies, as evidenced by the everyday news.

7. Acknowledgment

R. Boulet benefits from a post doctoral grant of the CARTAM–SAT project (FEDER 2007-2013 operational program in French Guyana). This study is partly funded (see Sec. 5) by the RTRA Science and Technology for Aeronautics and Space (http://www.fondationstae.net/) in Toulouse (France) under the MAELIA project (http://maelia1.wordpress.com/). The yEd Graph editor has been used for producing the figures. Statistical properties of networks have been computed with R and the library igraph (http://www.rproject.org/).

8. References


Nowadays, graph theory is an important analysis tool in mathematics and computer science. Because of the inherent simplicity of graph theory, it can be used to model many different physical and abstract systems such as transportation and communication networks, models for business administration, political science, and psychology and so on. The purpose of this book is not only to present the latest state and development tendencies of graph theory, but to bring the reader far enough along the way to enable him to embark on the research problems of his own. Taking into account the large amount of knowledge about graph theory and practice presented in the book, it has two major parts: theoretical researches and applications. The book is also intended for both graduate and postgraduate students in fields such as mathematics, computer science, system sciences, biology, engineering, cybernetics, and social sciences, and as a reference for software professionals and practitioners.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: