The network of the International Criminal Court decisions as a complex system

Fabien Tarissan[1] and Raphaëlle Nollez-Goldbach[2]

[1] Université Pierre et Marie Curie & CNRS, Laboratoire d'Informatique de Paris 6
[2] Université Paris Ouest, ENS & CNRS, Centre de Théorie et d'Analyse du Droit

Abstract. Many real-world networks lend themselves to the use of graphs for analysing and modeling their structure. This approach has proved to be very useful for a wide variety of networks stemming from very different fields. Yet, only few papers focused their attention on legal networks. This paper intends precisely to remedy this situation by analysing a major legal network by means of complex system methods. The network under investigation is the network composed by decisions taken by the International Criminal Court since its creation. We first model the network by a simple directed graph in which nodes are the decisions and links represent citations between decisions. Our analysis shows that standard properties shared by common real networks are also present in this network. Then we turn to studying the network by means of bipartite graphs that involve both decisions and articles of law. We show that this two-level structure presents several non trivial properties and we show evidences of the relevance of the bipartite representation to explain properties observed in the graph of citations.

Keywords: Complex networks, Bipartite graphs, Legal studies

1 Introduction

Many real-world networks lend themselves to the use of graphs for analysing and modeling their structure. We can cite among others actor networks [1, 2] which relate actors performing in the same movies or authoring networks [2, 3] which relate authors publishing together. Since the seminal paper of Watts and Strogatz [1], it has been shown that different kind of networks yet share similar non trivial properties, such as an heterogeneous distribution of their degrees, a high local density, short distances, etc.

Since then this approach has been widely used in many different fields, ranging from computer science (the Internet, peer-to-peer networks, the web), to biology (protein-protein interaction networks, gene regulation networks), social science (friendship networks, collaboration networks), linguistics, economy, etc. Thus, the complex system theme has opened a new interdisciplinary interplay between fields that were usually separated. Recently, law and computer science followed this promising approach, using methods related to complex systems to model and analyse legal networks [4, 5]. It is this rich and promising field of research this paper intends to commit itself to by studying the networks composed of the decisions taken by the International Criminal Court since its creation.

Following the traditional approaches, we use standard directed graphs to represent the networks. Here the nodes stand for the decisions and the links for citations between decisions. Although useful, such a simple representation is not particularly close to the real structure of the networks and does not account for the inherent complexity and hierarchy commonly observed in real data. In the context of legal networks in particular, interactions between decisions take place at various levels. Indeed, to motivate their decisions, judges usually rely on former decisions – first level exhibiting direct interactions – but also refer to articles of the ICC Statute and norms and regulations of the Court – second level pointing out indirect interactions. This induces a two-level structure in which direct and indirect relations interplay in the ruling process.

The existence of two-level structures in real networks led the community to use also *bipartite graphs* , i.e. graphs in which nodes can be divided into two disjoint sets, $\top$ (e.g. articles) and $\bot$ (e.g. decisions), such that every link connects a node in $\top$ to one in $\bot$. Bipartite graphs are fundamental objects which have proven to be very efficient for both the analysis [7, 9, 8] and the modeling [6, 10] of complex networks as they are able to reveal patterns that could not have been detected on simple graphs. The present study follows this approach and relies at the same time on simple direct graphs and bipartite graphs in order to extract all relevant properties from the network.

The remaining of the paper is organised as follow: Section 2 will review the technical background necessary for going throughout the paper; Section 3 will present the main results obtained and finally Section 4 will conclude the paper and open on new perspectives.

2 Background

In this section, we introduce the required background for the remainder of the paper. First, we focus on the dataset (Section 2.1) extracted from the ICC database. Then, we discuss the several frameworks we used in order analyse its structure (Section 2.2).

2.1 Legal networks

The International Criminal Court (referred to further as ICC) is the first permanent international criminal jurisdiction, established to judge international crimes (genocides, crimes against humanity and war crimes). The creation of the Court is very recent (2002) and only 18 cases are currently opened.

The main production of the Court are decisions, which are legal statement ruling on juridical issues. In order to motivate their decisions, judge may either rely on former decisions of the Court, or rely on articles of the ICC Statute, and/or norms and regulations of the Court. Here below is an example of such a motivation found on a footnote of decision ICC-01/04-01/06-2126-Anx:

APPEALS CHAMBER, JUDGEMENT ON THE APPEALS OF THE PROSECUTOR AND THE DEFENCE AGAINST TRIAL CHAMBER I'S DECISION ON VICTIMS' PARTICIPATION OF 18 JANUARY 2008, 11 JULY 2008, **ICC-01/04-01/06-1432**, PARA. 95. SEE ALSO TRIAL CHAMBER II, DECISION ON THE MODALITIES OF VICTIM PARTICIPATION AT TRIAL, 22 JANUARY 2010, **ICC-01/04-1/07-1788**, PARA. 30. SEE ALSO DEFENCE FOR GERMAIN KATANGA'S ADDITIONAL OBSERVATIONS ON VICTIMS' PARTICIPATION AND SCOPE THEREOF", 10 NOVEMBER 2009, **ICC-01/04-01/07-1618**: "IT HAS BEEN HELD THAT **article 69(3)** GIVES THE COURT A GENERAL ...

In this example, one can notice the two types of arguments used by the judges. The text clearly refers to former decisions (highlighted in red) but also refers to article of the ICC Statute (in blue).

In the rest of the paper, we only focus on the Lubanga case (DRC situation) to concentrate on the most advance case. It involves approximately 7 000 decisions and 1 500 articles.

2.2 Graph frameworks

Directed graphs. As depicted in the introduction, it is quite natural to represent the network as a directed graph $G = (V, E)$, with $n = |V|$ and $m = |E|$, in which the nodes stand for the decisions – identified by their ICC number – and a link between nodes u and v exists if the decision u cites the decision v. Note that technically, the graph is a Directed Acyclic Graph (DAG) since, for obvious reasons, the decisions can only refer to existing ones. Thus there is no cycle in the network.

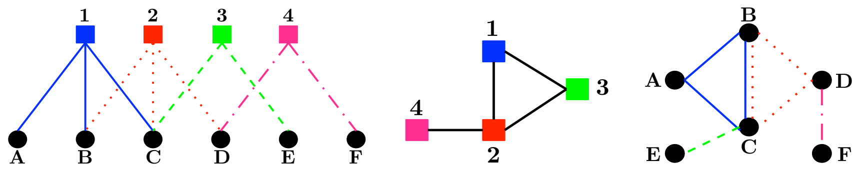
This defines the graph of citations among decisions, which will be referred to simply as the *graph of citations*. By doing so, we can compute standard metrics and compare the results to what is obtained on usual complex networks. According to standard studies, one usually observe for instance that graphs are sparse, *i.e.* the density $\delta = \frac{2.m}{n.(n-1)}$ is very small, and their degree distribution is heterogeneous (often close to a power-law).

Another key property concerns the local density which is meant to study how dense a neighbourhood of a node is in the graph. This concept is generally captured by the clustering coefficient $\texttt{cc}(G)$ or the transitivity ratio $\texttt{tr}(G)$ [1, 13, 14], defined formally by:

$$\texttt{cc}(G) = \frac{\sum_v \frac{\Delta(v)}{\vee(v)}}{n} \qquad\qquad \texttt{tr}(G) = \frac{\Delta(G)}{\vee(G)}$$

where, for each $v \in V$, $\Delta(v)$ denotes the number of directed triangles (sets of three nodes u, v, w, such that $(u, v), (u, w), (v, w) \in E$) which originate at v; $\vee(v) = \frac{d(v).(d(v)-1)}{2}$ denotes the number of pairs of neighbours of v which computes the number of *possible* directed triangle; $\Delta(G) = \sum_v \Delta(v)$; and $\vee(G) = \sum_v \vee(v)$. Note that the clustering coefficient of a node can be defined for the in-degree and the out-degree, that is when the node is at the origin of the directed triangles (case of u in the example above) and when it is at the end (case of w). Both variants makes sense and will be investigated in Section 3.

A classical observation in complex network studies is that all these quantities are high, at least compared to the density δ of the graph. Note however that



(a) bipartite graph (b) $\top$-proj. (c) $\bot$-proj.

Fig. 1. Example of bipartite graph and its $\{\top, \bot\}$-projections

the meaning of the existence of such a pattern depends on the context of the network. It has been shown that it could be related to robustness properties of the network, or properties related to dynamical aspects of the networks (see for instance [11, 12] for biological cases).

Bipartite graphs. As stated in the introduction, the previous formalism does not account for higher level of relations between the decisions. In particular, one does not exploit the references made to the articles which relate the decisions to the articles (and regulations and norms) they refer to. This two-level structure calls for a specific framework that is perfectly matched by the concept of bipartite graphs.

A *bipartite graph* is a triplet $G_b = (\top, \bot, E_b)$, where $\top$ is the set of *top* nodes (here the articles/norms/regulations), $\bot$ the set of *bottom* nodes (here the decisions), and $E_b \subseteq \top \times \bot$ the set of links that relate the decisions to the articles. We denote by $n_\top$ (resp. $n_\bot$) the number of top nodes (resp. bottom nodes).

Compared to standard graphs, nodes in a bipartite graph are in two disjoint sets, and the links are always between a node in one set and a node in the other set. An example of bipartite graph is given in Fig. 1(a), where $\top$ nodes are depicted by squares and $\bot$ nodes by circles.

The $\bot$-projection of G_b is the graph $G_\bot = (\bot, E_\bot)$ where two nodes (of $\bot$) are linked together if they have at least one neighbour in common (in $\top$) in G_v: $E_\bot = \{(u, v), \exists x \in \top : (u, x) \in E_b \text{ and } (v, x) \in E_b\}$. The $\top$-projection is defined dually. Both projections are illustrated in Fig. 1(b) and 1(c). Thus, in our case, the $\bot$-projection corresponds to a graph of decisions, such as G, but a link between two decisions exists if and only if there is at least one common article to which they both refer.

Note that by projecting a bipartite graph into a simple graph, we can then reuse all the metrics defined above for standard graph. But we can also compute specific metrics for bipartite graphs, such as $k_\top$ (resp. $k_\bot$) the average degree of top nodes (resp. bottom nodes) and $\delta_\mathsf{b} = \frac{m_\mathsf{b}}{n_\top \cdot n_\bot}$ the density of the bipartite graph.

	Directed graph	
	Complete	Connexe
n	6 894	1 575
m	17 625	3 319
δ	0.00	0.00
k	2.5	2.1
d^+	214 / 158	433 / 433
`tr`	0.03 / 0.10	0.03 / 0.08
`cc`	0.14 / 0.19	0.12 / 0.22

	Bipartite graph
$n_\top$	1 415
$n_\bot$	6 894
$m_\mathtt{b}$	11 371
$n_\bot(conn.)$	1 683
$k_\top$	1.7
$k_\bot$	8.0
$\delta_\mathtt{b}$	0.00
$d^+_\top$	802
$d^+_\bot$	116

Table 1. Global statistics for the graph of citations (left) and decision/article bipartite graph (right).

3 Results

The purpose of this section is to position the ICC decisions network as regard the properties observed in common complex networks. To do so, we start investigating global statistics (Section 3.1) before focusing on more specific properties such as the degree distribution, the local density and some correlations in the bipartite structure (Section 3.2).

3.1 A global perspective

The first statistics we focus on concern some basic properties observed in most real-world networks, formally presented in the previous section. Table 1 presents the results both for the graph of citations (left) and the bipartite graph (right).

As expected, all usual observations made on real-world networks stand also for the legal network under investigation. More precisely, one can see that the graph is sparse ($\delta = 7 \times 10^{-4}$) and that the local density (both the transitive ratio and the clustering coefficient) are several orders of magnitude higher.

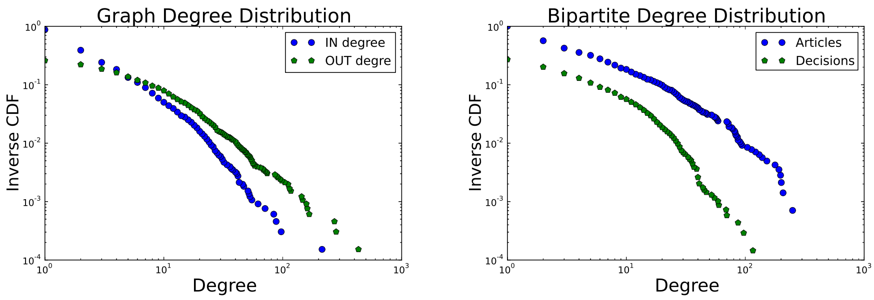
Beside one can notice that the highest degree d^+ is also way higher than the average degree k. This indicates some heterogeneity in the degree distribution that will be investigated further.

As regard the statistics on the bipartite graphs, it also matches the expectations since the graph is also sparse and both top and bottom highest degrees ($d^+_\top$ and $d^+_\bot$) are two orders of magnitude higher than their respective average degree ($k_\top$ and $k_\bot$).

Those results on the global structures of the two graphs are now refine in the following section.

3.2 A deeper analysis

Degree distribution. In order to refine the general statistics presented above, Figure 2 presents the degree distribution observed in the graph of citations (left)



(a) Directed graph

(b) Bipartite graph

Fig. 2. Inverse cumulative degree distributions for the graph of citations (left) and the bipartite graph (right).

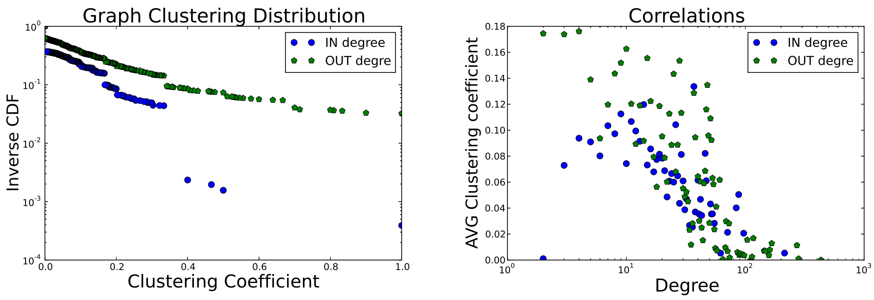
and the bipartite graph (right). In the two figures, the horizontal axis, in log-scale, stands for the degree of the nodes, while the vertical axis, also in log-scale, presents the inverse cumulative proportion (according to the total number of nodes). The shape of the plots, close to a straight line in the log-log scale over several orders of magnitude, thus confirms that we are dealing with heterogeneous distributions.

One can notice that for the out degree (the number of citations a decisions makes) in the graph of citations, the proportion is higher than the in-degrees (the number of references made to a decision). This is in particular true for the high degrees, which is well explained by the existence of annexes in the corpus. Indeed, those usually list all the decisions that a case has referred to.

Note also that, although the high out-degrees are not particularly meaningful since it corresponds to annexes, it is however particularly relevant for the in-degrees. Indeed, a manual investigation showed that the top-3 of cited decisions corresponds for major ruling in the case. The first one deals with the conviction of the accused, the second with its sentence and the last with the remedy and reparations of the victims, which are all three important issues in the international Court.

As regard the bipartite structure, one can also notice the same kind of comparison between the degrees of the articles and the ones of the decisions. This indicates that, from a global point of view, each article tends to be more referred to that each single decisions. This can be explained by the status of the articles of the ICC Statute towards decision. This remark is also corroborated by the highest degrees of articles and decisions (see Table 1 right).

Clustering coefficient. Figure 3 presents several properties related to the clustering coefficient, as defined in Section 2. The left part displays the inverse cumulative distribution (in lin-log scale) of the coefficients in the graph of citations



(a) Directed graph (b) Directed graph

Fig. 3. Clustering coefficient and correlations in the graph of citations

defined for the nodes that originates the directed triangles (out-degree, in green) and the ones that are at the end of the directed triangles (in-degree, in blue). Note that the plots are normalised over the number of nodes with degree ≥ 2 in order to avoid side effects from the nodes of degree 1, for which the notion of clustering coefficient is inadequate.

Interestingly, one can notice that the proportions of nodes that are at the origin of triangles is particularly high compared to the one that are at the end. This indicates that when a decision refers to two other decisions, one of them tends to cite the last one very often. Manual investigations performed on the triangles that involve highly cited decisions showed that the over-representation of this pattern is meaningful as regard legal networks. It turned out that most of those triangles involve a decision w that is contested by the accused, thus leading to a decision v that cite w. Then a final decision u rules between the issue, thus citing both v and w. This leads to the directed triangle originated by u with w as end-point.

The right part of Figure 3 presents a non trivial correlation between the clustering coefficient of a node and its degree in the graph of citations. More precisely, a dot (x, y) in this plot corresponds to the fact that nodes with degree x have, in average, a clustering coefficient of y. The figure shows that the higher the degree, the lower the clustering coefficient in average. This is true both for in-degree and out-degree except for very low degrees. Indeed, the case of degree-2 nodes is very different depending on whether we consider the origin of the triangles or the end. For the end, the coefficient is very low (close to 0), which shows that when two decisions cite a third one, they usually do not rely on each other (blue dots). On the other end, if one decision cite two different decisions, those two decisions tends to be related (green dots). The latter case is easily explained by the remark made in the previous paragraph since in the former example, the decision u has precisely degree 2 and its clustering coefficient is 1 thus increasing the average clustering coefficient of degree-2 nodes.

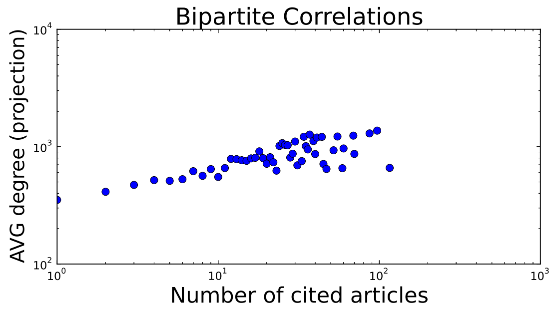


Fig. 4. Correlations

Bipartite/Projection correlations. Finally, Figure 4 presents a correlation between the degree of a decision in the bipartite graph and its degree in the projection. More precisely, a dot (x, y) in the plot stands for the fact that decisions that refer to x articles in the bipartite network are related (in the projection) to y different decisions in average.

The figure presents a strong correlation between those two quantities. The regular increase (in log-log scale) is natural since the more a decision refers to different articles, the more decisions it will be related to. What is less intuitive though, is why the slope of the progression is not higher. Indeed, the reader might have noticed that according to the plot, multiplying by 100 the number of references to articles only multiplies the degree in the projection by 10. This fact clearly indicates a strong overlapping between the articles. This observation makes sense since decisions that concerns a similar content (that is, deals with similar legal issues) tend to refer to the same articles to motivate their content.

4 Conclusion

In this paper, we studied a legal network composed of decisions taken by the International Criminal Court since its creation. The analysis made on the most advance case of the Court shows that it completely matches standards properties shared by real-world networks, thus confirming the relevance of the complex system approach towards legal networks.

Besides, we investigated more in depth some of the properties related to the local density and provided a first interpretation of the over-representation of directed triangles in such networks. By using different formal frameworks, we also exhibited new properties such as the overlapping of articles in the bipartite representation of the decision/article network.

However, the different analyses performed in this study have been made independently. Yet, the strong patterns identified here call for the definition of a unique framework able to integrate both direct relations between decisions (such as the citation process) and indirect relations (such as the decision-article relations). This would entitle to consider correlations between those two levels of interaction that might shed light on important properties of the network. We let this promising approach for further studies.

Another appealing aspect which has not been investigated for the moment concerns the time. As stated in Section 2, the graph is actually a DAG since temporal aspects prevent from the existence of cycles in the network. This remark lead to consider the temporal evolution of the network instead of considering the whole decision network since the creation of the Court. We also let this aspect for future studies.

References

1. D. Watts and S. Strogatz, *Collective dynamics of small-world networks*, Nature, vol. 393, pp. 440–442, June 1998.
2. M. Newman, S. Strogatz and D. Watts, *Random graphs with arbitrary degree distributions and their applications*, Physics Reviews E, vol. 64, 2001.
3. M. Newman, D. Watts and S. Strogatz, *Random graph models of social networks*, PNAS, vol. 99, 2566–2572, 2002.
4. D. Bourcier and P. Mazzega, *Codification law article and graphs*, Legal Knowledge and Information Systems, JURIX 2007, A.R. Lodder and L. Mommers (eds.), IOS Press, 29–38, 2007.
5. R. Boulet and P. Mazzega and D. Bourcier, *A Network Approach to the French System of Legal codes - Part I: Analysis of a Dense Network*, Artificial Intelligence and Law, 19, 333–355, 2011.x
6. J-L. Guillaume and M. Latapy, *Bipartite graphs as Models of Complex Networks*, Physica A 371, pages 795–813, 2004.
7. M. Latapy, C. Magnien and N. Del Vecchio, *Basic notions for the analysis of large two-mode networks*, Social Networks, vol. 30, no. 1, pp. 31–48, January 2008.
8. Y.-Y. Ahn, S. E. Ahnert, J. P. Bagrow and A.-L. Barabási, *Flavor network and the principles of food pairing*, Nature Scientific Reports, 2011.
9. M. Tumminello, S. Miccichè, F. Lillo, J. Piilo and R. Mantegna, *Statistically Validated Networks in Bipartite Complex Systems*, PLoS ONE, vol. 6, 2011.
10. F. Tarissan, B. Quoitin, P. Mérindol, B. Donnet, M. Latapy and J-J. Pansiot, *Towards a Bipartite Graph Modeling of the Internet Topology*, Computer Networks, vol. 57, no. 11, pp. 2331–2347, Elsevier, 2013.
11. S. Mangan, S. Itzkovitz, A. Zaslaver and U. Alon, *The Incoherent Feed-forward Loop Accelerates the Response-time of the* gal *System of* Escherichia coli, JMB, Vol. 356, pp. 1073–1081, 2006.
12. S. Mangan and U. Alon, *Structure and function of the feed-forward loop network motif*, PNAS, vol. 100, pp.11980–11985, 2003.
13. T. Schank and D. Wagner, *Finding, counting and listing all triangles in large graphs, an experimental study*, in Workshop on Experimental and Efficient Algorithms (WEA), 2005.
14. T. Schank and D. Wagner, *Approximating clustering coefficient and transitivity*, Journal of Graph Algorithms and Applications (JGAA), vol. 9:2, pp. 265–275, 2005.