Temporal networks of 'Contrafacta' in the first three troubadour generations

Stefano Milonia
Institute of Advanced Studies, University of Warwick, Coventry, UK

Matteo Mazzamurro
Department of Computer Science, University of Warwick, Coventry, UK

Abstract

The scope of this research is that of revealing the interconnected nature of medieval Romance lyric by looking at musical imitations. In the Middle Ages, melodic imitation was an essential part of artistic creation as old melodies were constantly borrowed by new authors, a practice known as contrafaction. In this study, we propose to analyse the complex relations between medieval lyricists resulting from this practice using temporal networks. We construct networks by representing each author's body of work as a single node and connecting a lyricist' corpus to that of another lyricist via a directed link when the latter author borrowed a melody from the former. To each directed link, we associate the temporal information of when the imitation was composed. Such networks provide a convenient visualization tool to explore the dataset and its connections in an intuitive fashion and are available online at: https://medmus.warwick.ac.uk/networks. They also provide an analytical tool: we use the networks to show how ideas might have spread among lyricists along time-respecting paths, and obtain measures of the authors' centrality and influence on the overall literary corpus. We compare the results obtained with the temporal networks with those obtained via more traditional centrality measures computed for corresponding static networks, and explain why the temporally informed measures may provide a more accurate depiction of authors' influence.

Correspondence:
Stefano Milonia, Institute of Advanced Studies, University of Warwick, Zeeman Building, Coventry CV4 7AL, UK.
E-mail:
stefano.milonia@gmail.com

1 Introduction

The texts that form the corpus of medieval Romance lyric cannot be considered independently from each other: the dialogues between authors, their relationships with their patrons, the events and political issues they refer to effectively make these literary works nodes of a complex system of relationships.

The objective of this paper is to propose efficient ways to visualize and analyse networks that synthesize a large amount of data pertaining to literary texts, namely metrical and musical imitations, which allow us to establish the existence of relationships between authors and between their works. These networks summarize the influence and receptivity of medieval poets and emphasize how their relationships can be analysed through multiple notions of centrality, while using temporal coordinates to draw chronologically coherent paths of influence. This study will allow scholars to draw a more accurate evaluation of

https://doi.org/10.1093/llc/fqac018

authors' importance within their literary tradition and thus rethink established literary canons.

This research is part of the project 'Connecting Medieval Music', which focuses on relationships between authors of medieval lyric in France, Occitania, Italy, and the Iberian Peninsula. Gathering and cross-examining archival, literary, and musicological data, the project aims at drawing a wide picture of the flows of culture and exchanges between centres of artistic production in medieval Europe, by focusing on the key figures that contributed to its development. If the analysis of information concerning the political and economic domain determines the historiographical picture of a territory in its more traditional sense, the study of relationships established by the practice of melodic imitation can reveal a layer of cultural history hitherto unexplored in a systematic manner.

The analysis of musical imitations requires a thorough examination of the single melodies and reveals intertextual connections that cannot be easily integrated in a homogeneous dataset. However, medieval lyric allows us to catalogue structural connections which can be demonstrated with a high degree of certainty thanks to the widespread practice of 'contrafaction', a technique that consists of reusing a pre-existing melody for composing a new text.

This practice calls for an intertextual perspective intended in its broader sense, as it does not only involve texts that consciously enter into a dialogue with their models, but it also establishes more subtle or indirect mnemonic relationships that contribute to the creation of an artistic environment. Within this environment, individual texts are the result of multiple stimuli which direct the development of thematic and stylistic aesthetics as well as practical skills, thus influencing textual content, music composition, and versification techniques.

The affirmation of romance lyric poetry is documented from the beginning of the 12th century and it is an inherently musical artistic form. Versification—intended as the distribution of the text in syllabic schemes recurring in every stanza—is correlated to, or arguably determined by, the fact that these texts were intended to be sung. Indeed, in each stanza, every syllable corresponds to a note or a group of notes; therefore, in a stanza of a song, the number of syllables always coincides with the number of melodic units. By virtue of this structural correspondence, the author was made certain that by iterating a same syllabic structure in every stanza these could be sung on the same tune. For the same reason, manuscript sources provided with music (with the exception of peculiar cases and non-strophic compositions such as *lais* and *descortz*) present the melodic notation only above the lines of the first stanza.

This process of composition suggests that authors were skilled in adapting a text to a determined melody, a custom inherited from widely spread practices in the liturgic, paraliturgical, and profane Latin repertory. The creation of 'contrafacta' was thus rather simple: the author of the new composition could easily write a text with the same syllabic structure of its model (either by actively counting syllables or instinctively replicating the prosody in signing) so as to more easily tailor the new text to the pre-existing melody. Since many lyric works that are extant in manuscript sources are not provided with musical notation, making a direct comparison of the melody impossible, it is mostly thanks to the repertories of metrical schemes (Frank, 1953–1967; Tavani, 1967; Mölk and Wolfzettel, 1972; Solimena, 1980; Antonelli, 1984; Solimena, 2000) that we are able to identify a possible relationship between models and contrafacta. However, metrical correspondence alone does not necessarily imply a melodic reuse, and, where the music has not been preserved, the identification of contrafaction has to be made with caution: the probability that two or more texts sharing the same syllabic scheme are linked to each other is proportional to the complexity and rarity of the syllabic scheme itself (Billy, 2002). Furthermore, the rhyme scheme constitutes a significant element to be considered in the identification of such relationships. While the disposition of the rhymes within the stanza does not directly affect the musical performance, in most instances the author of the contrafactum adopts the rhyme scheme of its model, often reusing the same rhyme sounds, and on rare occasions even rhyme-words. The cross-examination of this information allows us to gather reliable data on the connections between authors and to establish, with some exceptions, the direction of the relationship between model and imitation.

The domain of contrafacta has been studied both by musicologists (Gennrich, 1956; Gennrich, 1965; Rossell, 2002) and Romance philologists (Chambers, 1952; Marshall, 1980; Billy *et al.*, 2003 ; Schultze, 2004; Lannutti, 2008; Milonia, 2016; Persico, 2017) in the poetry of all European vernacular languages. Since metrical and musical features can be easily reused in different linguistic contexts, studies on contrafacta paid particular attention to the influence of a lyric tradition on authors from a different geographical and linguistic setting (Frank, 1952; Schultze, 1989; Asperti, 1991). A visualization of a network of troubadour contrafacta has also been realized by Chapman (2020). From our perspective, the analysis of a network based on connections established by metrical–musical elements is particularly useful in order to allow a more complete vision of medieval Europe's cultural history. This is because the practice of contrafaction is strictly linked to lyric genres that traditionally provide more contextual information, allowing one to connect artistic work with the geographical and historical environment that engendered it. Medieval Occitan treatises on poetic versification (*Leys d'amors*) legitimize the reuse of a melody for specific genres and forbid it for others. If love songs are to be composed on an original tune, political songs (*sirventes*) or laments (*planhs*) can be, and more often are, contrafacta; consequently, texts with a borrowed melody tend to contain more references to people, places, and events.

2 Methodology

2.1 Networks theory

Networks provide a versatile tool to represent and analyse sets of entities in which some pairs are related to one another and are thus ubiquitous in the study of a variety of real-world systems, from the connections in social networks (Wasserman and Faust, 1994), to citations in the scholarly literature (Zhao and Strotmann, 2015) and in patents (Batagelj *et al.*, 2014: 175–215). The entities in a network are called nodes and the relations between any two nodes are called links.

In this study, we will be interested more specifically in temporal networks (Holme and Saramäki, 2019), occasionally also referred to as time-varying graphs (Whitbeck, 2012) or evolving graphs (Xuan *et al.*, 2003). As opposed to standard static networks, in which links are 'fixed', in a temporal network links are transient and exist only in certain intervals of time. As such, a temporal network is perhaps easier to imagine as a sequence of static snapshots of the network taken at different moments in time, each capturing a specific instance of the network's evolution, a perspective which reconnects them to the more encompassing theory of multilayer networks (Boccaletti *et al.*, 2014). Temporal networks have been successfully used, among other scopes, to model email exchanges and the consequent diffusion of information: in Holme (2005), each individual is represented as a node in a network and an email is a directed link between two individuals that exists for a single instant in the moment it is read. Similarly, temporal networks are a crucial tool in the study of the spread of diseases: in Riolo *et al.* (2001), individuals are again taken as nodes, and if two individuals come in proximity of each other, a link is drawn between them which persists for the duration of their contact; transmission of disease may then happen from an infected individual to a healthy one along such links. Temporal networks are also used as a tool to analyse the emergence of patterns in static networks (Medo *et al.*, 2011).

2.2 Node centrality in static networks

Although the majority of real-world networks are evolving in time, for simplicity it is sometimes convenient to 'ignore' the temporal information and consider the static network including all links appearing in the temporal network at any point in time. Such a cumulated view of the network is called the temporal network's projection (Holme and Saramäki, 2019). Several measures exist to analyse the importance of nodes in static networks.

The 'degree' of a node is the number of its connections and thus arguably the simplest notion of importance of a node. For a directed network, one distinguishes between the out-degree of a node, that is the number of connections issuing from the node, and its in-degree, that is the number of connections ending in the node.

A path from a node v to a node v' in a directed network is a sequence of nodes connected by directed links that starts in v and ends in v', with the arrows never pointing back to an already visited node. If a node v is connected via a path to a node v', we say that

v' is reachable from v. The set of all nodes reachable from a node v is called the 'reachability set' of v. The size of the reachability set extends the notion of (out-)degree and can be thought of a measure of the potential spread of information released from v when this can be transmitted from v to the nodes it is connected to it and from these to the nodes they connect to, and so on. The size of the reachability set can be a more appropriate notion of importance of a node in the context of information diffusion since, while a node may have few connections in itself, the connections of these may be numerous and the information it released may thus spread extensively.

As many possible paths may exist between any two nodes in a network, special attention is given to the shortest paths connecting two nodes, by which here we simply mean the paths containing the smallest number of links. Shortest paths are crucial to the definition of two other measures of importance of a node: the closeness and the betweenness centrality.

'Closeness centrality' can be seen as a less trivial extension of the out-degree, incorporating information on how far two nodes are from each other. According to this measure, a node is more 'central' when it is at a close distance to the other nodes, as measured by shortest paths. The formal definition of closeness centrality as originally proposed in Beauchamp (1965) and reprised in Freeman (1978) and Latora and Marchiori (2007) is problematic when not all nodes of the networks are connected by paths, which is often the case for real-world directed networks. Therefore, we will follow here its generalization to potentially not connected networks, explicitly given in Marchiori and Latora (2000), but previously already mentioned in Gil-Mendieta and Schmidt (1996). Such definition of closeness centrality is also occasionally referred to as harmonic centrality. Consider again a node v. For each node $u \neq v$, let d_{uv} be the distance from u to v, by which we mean the minimum number of steps needed to reach v from u. If no path connects u to v, formally let $d_{uv} = \infty$ and $\frac{1}{d_{uv}} = 0$. Similarly, let d_{vu} be the distance from v to u, remarking that, in a directed network, d_{uv} and d_{vu} may be different. Then the closeness (harmonic) centrality of v is defined, for incoming paths, as the sum of the reciprocal of the distances from all nodes other than v to v, that is

$$C_C^{\text{in}}(v) = \sum_{u:u\neq v} \frac{1}{d_{uv}}.$$

Equation (1): Closeness centrality for incoming paths

Similarly, for outgoing paths, the closeness centrality is defined as the sum of the reciprocal of the distances from v to all nodes other than v, that is

$$C_C^{\text{out}}(v) = \sum_{u:u\neq v} \frac{1}{d_{vu}}.$$

Equation (2): Closeness centrality for outgoing paths

This implies that if v can be reached from the other nodes in few steps, then its centrality $C_C^{\text{in}}(v)$ is large, while if it is far from them, its centrality $C_C^{\text{in}}(v)$ is small. Similarly, if from v one can reach the other nodes in few steps, then $C_C^{\text{out}}(v)$ is small, otherwise $C_C^{\text{out}}(v)$ is large.

'Betweenness centrality' is based on the principle that a node v is more central when it is 'strategically located' along the largest number of shortest paths connecting other nodes (Bavelas, 1948; Freeman 1977). More formally, if for any two distinct nodes u, w, with $u \neq v \neq w$, we let g_{uw} be the number of shortest paths connecting u to w and we let $g_{uw}(v)$ be the number of such paths passing through v, then the betweenness centrality $C_B(v)$ of v is defined as

$$C_B(v) = \sum_{\substack{u,\,w \\ u\neq v\neq w}} \frac{g_{uw}(v)}{g_{uw}},$$

Equation (3): Betweenness centrality

that is, we are summing over all pairs of distinct nodes the proportion of shortest paths connecting them that passes through v.

Each of the above measures represents a different interpretation of the importance and potential influence of a node in the overall network. The classic (Freeman, 1978) contains a very clear discussion of the meaning and origin of centrality measures, while Latora and Marchiori (2007) provides a more modern and extensive account.

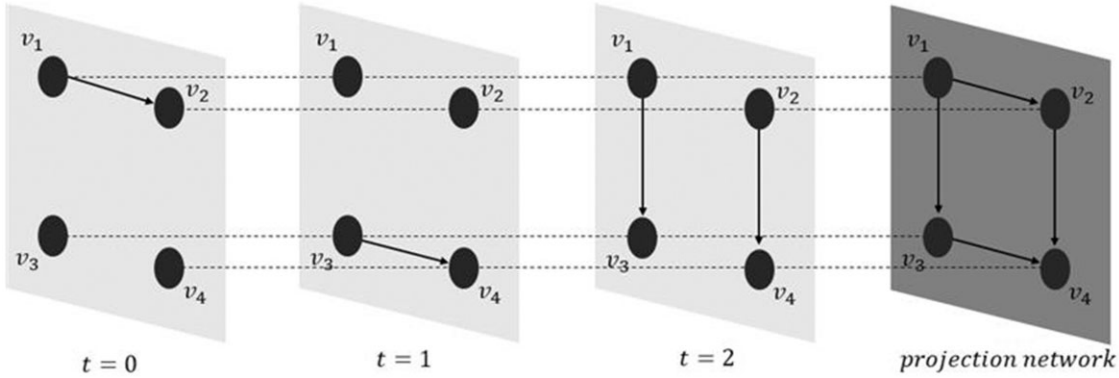

Fig. 1. Simple temporal network with its projection

2.3 Node centrality in temporal networks

The definitions given above for static networks need some care to be generalized and correctly applied to the analysis of the projection of a temporal network.

First of all, it is important to notice that not all paths that appear in the projection network are viable, or time-respecting paths (Holme and Saramäki, 2019) and this is crucial in the study of the transmission of information from a node to another one. To clarify this statement, let us consider the situation presented in Fig. 1. In the projection network, two paths connect v_1 to v_4: one passing through v_2 and the other one passing through v_3. The link connecting v_1 to v_2 appears at time $t = 0$ and the link connecting v_2 to v_4 appears at time $t = 2$. Thus, if the links represent transmission of information, a piece of information from v_1 would be able to reach v_4 by first being transmitted from v_1 to v_2 at time $t = 0$ and then from v_2 to v_4 at $t = 2$. On the other hand, the link connecting v_1 to v_3 appears at time $t = 2$, while the link connecting v_3 to v_4 appears at $t = 1$. This means that it would not be possible for information from v_1 to reach v_4 via v_3. Hence, the only time-respecting path connecting v_1 to v_4 is the one passing through v_2.

Second, in a temporal network, the notion of shortest path (among the time-respecting ones) may have two different interpretations and thus impact centrality measures. On the one hand, a shortest path connecting an ordered pair of nodes may be, as in a static network, a path that allows to reach the latter from the former in the smallest number of steps; on the other hand, a shortest path may be one that allows to reach the latter from the former in the shortest time (Vega and Magnani, 2019). Which of these two interpretations should be preferred depends on the context. In our case, as explained in later sections about the construction of the networks, we will make use of the number of steps in our definition of shortest paths and thus closeness and betweenness centrality. Nicosia *et al.* (2013) and Taylor *et al.* (2019) define several careful alternative generalizations of centrality measures to temporal networks and Tang *et al.* (2013) presents their applications to real-world networks.

3 Our Experiment

3.1 Dataset

To study the practice of contrafaction and the connections it weaved among medieval troubadours, we used the 'Connecting Medieval Music' database: the data pertaining to the Occitan repertoire gathered within relies on information from the 'Bibliografia Elettronica dei Trovatori' (BEdT: Asperti and de Nigro, 2012), a relational database of Provençal troubadour texts, and the 'Dizionario biografico dei trovatori' (Guida and Larghi, 2014). It is further complemented by critical editions and secondary literature. The database provides schematic information on each text, including its author and its creation date, most frequently given as a year range, when the exact composition year is not available. Furthermore, the database specifies

whether the song's metric scheme was directly borrowed from another text (i.e. if it is a 'contrafactum'), and, more generally, if metrical analogies exist with other compositions. Based on their period of activity, BEdT subdivides authors into six roughly chronological generations. In this study, we considered all songs written by authors of the first, second, and third generation, composed for their greatest part between the beginning of the 12th century and the first decades of the 13th century. Among these, we selected those which either served as models for contrafacta or presented metrical analogies with other songs. The difference between these two groups does not lie in the preservation of the music for both model and contrafacta, which, as mentioned, is an exceedingly rare occurrence in the Occitan repertoire. Rather, the latter category comprises those works whose relationship with its model cannot be demonstrated with sufficient confidence and those for which the direction of imitation is uncertain. Because these contrafacta and songs with metrical analogies were written by authors of all generations (even beyond the third), we were able to study the lasting influence and impact of texts composed by authors of the first three generations on the whole corpus of Provençal Troubadour songs. Models and contrafacta from other repertoires (namely French, German, and Latin) that presented links with Occitan authors of the first three generations were also included.

Medieval manuscripts constitute late witnesses of what is a small part of the actual lyric production, since most Occitan songbooks were composed between the end of the 13th and the beginning of the 14th century (Avalle and Leonardi, 1993) and a non-quantifiable amount of works are lost: it is thus important to acknowledge that reconstructing a network which accurately reflects the actual system of relationships connecting Occitan poets in the Middle Ages is only possible to a certain extent.

3.2 Data cleaning

As we mentioned above, the construction of temporal networks requires that each link be provided with precise temporal information about the instant it comes into existence and the moment it disappears. In a large number of instances, temporal indication about the composition of a song or an author's activity period was vague (e.g. 'early-13th century' or 'middle of the 13th century'), or hypothetical (e.g. 'after 1194, and before 1198, around 1196'), or completely absent. These cases called for the adoption of reasonable conventions, which we state here as schematically as possible:

> If the indications were vague, we used arbitrary but sensible and consistent numeric translation, so that 'Beginning of the 13th century' became 1201–1210, 'End of the 13th century' became 1291–1300, 'Between the 12th and the 13th century' became 1191–1220, and 'Middle of the 13th century' became 1226–1275.
>
> If there were multiple hypotheses about the composition date, we used the most inclusive range (e.g. for 'after 1194, and before 1198, around 1196' we used the range 1194–1198).
>
> If the composition date of a song was absent, we used the activity period of its author, as reported by the 'Dizionario Biografico dei Trovatori' (Guida and Larghi, 2014), although this may also be an approximate time span.
>
> If a song s_2 is a contrafactum of another song s_1, then s_2 must have been composed after s_1. Suppose for both songs a time range was given in the data or could be retrieved as described by the previous steps. Then the lower end of the time range of s_2 was tweaked to be either its current value or the lower end of the time range of s_1, whichever was larger.

Applying these conventions allowed to determine for each song in the dataset at the very least a broad indication of the time interval in which it was composed, and for each author an approximate activity period, in a way which made the temporal data self-consistent. This information is used for the construction of the networks as detailed below.

3.3 Networks construction

Using the above dataset and conventions, we constructed two networks and used them for visualization and analysis purposes. One network represents the connections between the songs of troubadours, as defined by contrafaction, and the other one is the temporal network of the relations the songs interwove between their authors.

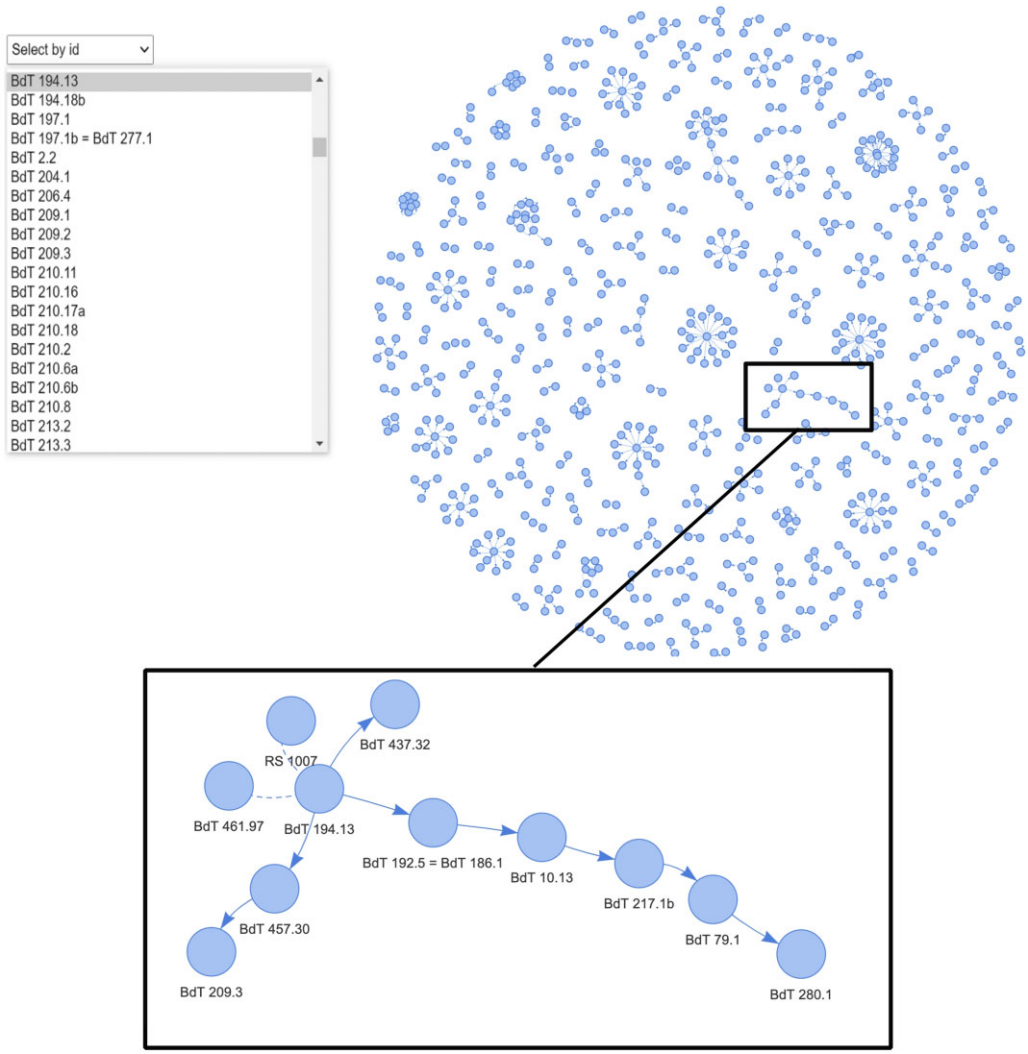

Fig. 2. Snapshot of the interactive widget of the network of songs, which reveals the existence of many components. The zoomed-in view shows the component with the largest diameter. Model BdT 194.13 ('L'autre jorn cost'una via', composed by Gui d'Uisel, active between 1195–1209), proved popular in subsequent years, inspiring the coblas BdT 437.32 ('Si com estau tain qu'esteja' by Sordel) and BdT 457.30 ('Physica et astronomia' by Uc de Saint Circ), to which responds Guillem del Baus' BdT 209.3 ('Liautatz ses tricharia'), and the tenso BdT 192.5 = 186.1 ('Seigner coms, saber volria', composed in 1215 by Gui da Cavaillo and Raimon VII de Toloza), which inspired the exchange of coblas initiated in 1220 by BdT 217.1b ('Bertran d'Aurel, si moria' by Guillem Figueira), joined by BdT 10.13 ('Bertram d'Aurel, s'aucizia', by Aimeric de Peguillan), to which Bertran d'Aurel responds with BdT 79.1 ('N'Aimeric, laissar poiria'); Lambert finally contributes to the dialogue with BdT 280.1 ('Seigner, cel qui la putia'). BdT 194.13 presents further analogies with the anonymous BdT 461.97 ('Dompna, s'ieu vos clamei amia') and with the French RS 1007 ('Comment que d'amours me deuille' by Blondel de Nesle)

3.3.1 *Network of songs*

In the network of songs (Fig. 2), each node corresponds to a song and is connected via a solid directed link to another song if the latter is a contrafactum of the former. If instead two songs present some more general metric analogies, the corresponding nodes are connected with each other via bidirectional dashed links, to denote the uncertainty of the directionality

of their relation. This simple network is used to visualize the influence of a model: the larger the out-degree or reachability set of a node, the larger the popularity of the corresponding song as a model. For example, the most popular song in the dataset is BdT 30.16 'La grans beutatz e.l fis enseignamens' by Arnaut de Maroill, whose reachability set contains eighteen songs, all contrafacta of the original model. As can be seen in Fig. 2, the network of songs is formed by a large number of small components, that is small sets of nodes connected with each other but not connected to any other node. Indeed, slightly less than a half of the components are pairs of songs and 75% contains at most four songs. Furthermore, most components have very small diameter, that is the longest directed path connecting two of its nodes is in general very short. The component with the largest diameter is shown in Fig. 2 and has diameter 5. These observations mean that the practice of composing a contrafactum of a song that is itself a contrafactum is relatively rare, mostly restricted to texts that are in direct dialogue with each other, where it is customary to use the same metrical structure and melody. This may be once again a consequence of our lack of information: faced with a group of lyrics that share a same metrical structure, scholars tend to identify one model for the entire group by looking at genre and chronology. However, we may not possess sufficient instruments to recognize that certain lyrics might be linked to the model indirectly, through one or more intermediary texts.

3.3.2 *Network of authors*

In the temporal network of authors, we let each node correspond to an author's body of work. Such a node comes into existence in the first year of the author's reported activity period and continues to exist from then on (the fact that an author stops composing after a given point in time does not imply that their body of work ceases to exist). We draw a solid directed link from an author's oeuvre to another's for each contrafactum the latter wrote using as a model a song composed by the former. Furthermore, we connect two authors' corpora using dashed directed links whenever two of their songs have metrical analogies. A directed link from an author's oeuvre to another's corresponds to the 'act of writing' of an imitation by the latter author using as model a song written by the former. This act is, in theory, instantaneous, and as such it should be associated with a specific instant of time t. Yet, as we have seen in the previous section, the temporal information about the composition of a song is not provided as an exact date in the database, but more generally as a year range, say $[t_1, t_2]$, where t_1 is the earliest possible year of composition and t_2 is the latest. What is known is that t lies at some point in such range. One thus needs to make a choice on how to fix a time t in this interval. We opted for fixing t as the year at the lower end of the interval, that is $t = t_1$. Other possible choices we considered are the end of the interval, its mid-point, and a random point in the interval, sampled uniformly. Each of these choices is arbitrary and can theoretically have an impact on the resulting temporal network. Experimentally, the impact was found to be occasionally remarkable on some nodes, but minor on the network overall (see also 'Supplementary Information' and 'Remarks and Limitations' for details). This proves that the inherent approximation in the dating of medieval lyrics does not affect the overall accuracy of the proposed models, strengthening the case for the use of temporal networks to analyse the dataset.

The projection of the temporal network of authors is used to visualize the potential influence of an author in terms of their role in the spread of information. The underlying assumption here is that authors who have imitated songs written by others might have, in some sense, internalized some of their stylistic features, and thus their subsequent composition might be signed by other authors' influence. In a way, this corresponds to the assumption that one's style and knowledge are influenced, directly or indirectly, by every source one actively engages with. Thus, a style propagates beyond the mere imitation and is reflected in an author's body of work, and thus in the work of those who imitate their works, and so on. To each connection in our network, we associate the BdT identifier of the model and of the contrafactum/imitation, as well as their composition dates. We display the network and the associated information in an interactive widget, which offers a high degree of personalization. The user can search the name of an author to highlight the corresponding node, together with all its adjacent nodes or all nodes in its reachability set, thus allowing an intuitive exploration of the authors in the dataset and their potential influence. The user can also choose

to display the dashed connections corresponding to metrical analogies or show only the solid connections corresponding to attested contrafacta. Beyond focusing on a single author in the network, one can choose to use the size and the colour of all the nodes to display measures of the influence of the corresponding authors, thus obtaining a more holistic view on the network. Such measures can be chosen among out-degree, in-degree, size of reachability set, closeness centrality, and betweenness centrality, so to more easily identify authors of great influence or receptivity according to the various criteria described in the 'Node Centrality' sections. Some examples of the visuals produced by the widget are shown in Section 3.5 ('Early Results and Examples'). The R code to generate, visualize, and analyse the networks can be found at https://github.com/MatteoMazzamurro/contrafacta-networks/releases/tag/v1.1.0 (Mazzamurro, 2022).

3.4 Remarks and limitations

A few technical remarks are needed in order to appreciate the properties and limitations of the above network models and measures, and their computation.

For the construction of the temporal network, we used the R package tsna (Bender-deMoll and Morris, 2016). This represents, as far as we are aware, one of the most advanced options available for the constructions and analysis of temporal networks. Yet, the package does not allow for a temporal network to be a multigraph at any time t. Thus if on year t, an author composed more than one contrafactum of another author's compositions, in the network we would connect the body of work of the latter to that of the former via a single directed link. In other words, the software does not distinguish between whether an author has composed one or more contrafacta of compositions from the same author in the same year. This might potentially affect some measures such as degree and betweenness centrality. However, given the rare occurrence of such event, the effects are very limited. Furthermore, closeness centrality and the size of the reachability set are unaffected by the presence of repeated imitations between the same pair of authors in the same time unit. Therefore, the network model and computation are still valid for our scope.

Temporal networks allow us to integrate together the temporal and the structural aspects of the connections between the nodes. In line with the main research aim of analysing the interconnectedness of troubadours through their lyrics, here we focus on the structural and topological aspects of the networks, using the temporal information to give a more meaningful definition to these features. We discuss concisely the temporal properties of the database in the Supplementary Information.

Another aspect worth noting is that some of the songs in the database are, in fact, 'tensos' (a dialogic composition in which different poets respond to each other in alternating stanza), and as such they have two or more authors. According to the construction of our networks of authors, if an author writes a contrafactum of a 'tenso', there will be links connecting each author of the tenso to the author of the contrafactum. Similarly, if the contrafactum of a song is a tenso, and thus is written by several authors, then several links will connect the author of the model to the authors of the contrafactum.[1]

It is thus worth noticing that there is no one-to-one correspondence between the links in our networks and contrafacta, as a single contrafactum may correspond to several links in the network. This aspect of the database does not constitute an issue, as we are focusing on the nodal properties of the network (i.e. the properties of the authors rather than the properties of the connections themselves), but would make it considerably more complicated to apply a link-based analysis to this network to explore, for instance, the role of a single contrafactum in creating connections between the authors, and users of our tool should be aware of this. Similar issues would emerge from any network construction centred on authors in this dataset, because the presence of tensos makes the act of writing a contrafactum a relation which is not necessarily binary. A possible solution could lie in the use of simplicial complexes (Atkin, 1977) to represent the connections between the authors, but this approach would considerably complicate the framework and goes beyond the scope of the current project.

A final remark concerns the length of a path connecting two authors, crucially in the definition of centrality measures. In this work, we preferred to use the number of steps in the path rather than the time elapsed from the composition of the first contrafactum in the path to the composition of the last one. This is because we believe that if an author has

imitated another author's song long after its original composition, and the imitation has proved very popular, this should be reflected in a larger betweenness centrality, and the time elapsed should not negatively affect the author's centrality.

3.5 Early results and examples

In this section, we show some examples of the visuals produced with the above-described widget, displaying in each example the results obtained from the computation of different centrality measures.

In Fig. 3, the size of each node represents its out-degree centrality, with larger nodes having larger out-degree, while the colour represents their in-degree centrality with lighter shades of grey corresponding to larger values of the nodes' in-degree. This provides a simple visual representation of how influential an author's body of work has been in the network in terms of how many songs their works have inspired (out-degree), and how many contrafacta and lyrics with metrical analogies the author wrote (in-degree). The lightest node at the centre of the whole network represents Peire Cardenal as the most receptive imitator of authors from the first to the third generation. The bottom window in the plot shows what happens when one selects an author in the widget and zooms in to read the details of their connections. In this case, we chose the author Guillem de Berguedan and highlighted his adjacent nodes for greater clarity. We notice that Guillem de Berguedan and Raimbaut III d'Aurenga appear to have similar sizes, and thus a similar degree of influence in terms of number of imitations of their works written by others, while Peirol and Bernart de Ventadorn are the most influential in terms of contrafacta inspired by their works and songs that can be connected to them by virtue of metrical analogies. Furthermore, Bertran de Born is qualified by a lighter colour, indicating that he has produced a greater number of imitations: he is indeed the most notorious composer of 'sirventes' in the troubadour tradition and the main author of contrafacta among this selected group.

Figure 4 shows a network of authors connected only by contrafacta, while metrical analogies have been excluded: the size of each node represents as above its out-degree, but in this case the colour represents the size of their time-respecting reachability set. In this case, the lighter shade of Folquet de Marseilla reveals that he had a potentially larger impact on the overall network, as his work has indirectly reached a large number of authors, even if he is similar to Bertran de Born in terms of sheer number of contrafacta inspired by their works. Note also that Folquet has the largest reachability set in this group (as shown by its lightest shade), in spite of the fact that Raimbaut de Vaqueiras' songs have proved very popular for imitations (as shown by its large size). This is due to the fact that authors who have imitated Raimbaut's works further down the line were probably relatively minor authors compared to those who imitated Folquet's works.

Similar detailed analysis concerning the betweenness centrality of the authors and/or their closeness centrality are possible (see the Supplementary Information for the corresponding networks). For both measures, Peirol emerges as a central figure, both in virtue of his acting as a 'bridge' towards otherwise poorly connected nodes—thus constituting the only available link for reaching many otherwise isolated poets—and because of having inspired a large number of authors, some of whom (such as Raimon de Miraval) were rather influential themselves.

If one is interested in comparing two node-based measures and is not immediately interested in displaying the connection between the nodes themselves, a simpler Cartesian plot can be used. In the two Cartesian plots in Fig. 5, for example, each dot corresponds to an author's body of work. Its position on the horizontal axis denotes the number of nodes in his reachability set, and hence points to the right are potentially more influential in the network in terms of their direct and indirect information transmission. Its position on the vertical axis denotes the number of authors who might have influenced him, and hence the points in the upper part of the graph represent those whose work has directly or indirectly collected the ideas and styles of a larger number of authors. In the top plot, we show the results obtained when we consider connections determined by writing any kind of imitation of another author's work. In the bottom plot, we consider the connections corresponding to contrafacta only. When we consider also metrical analogies, temporal aspects become soon more apparent. On the top left corner of the plot, we find dots corresponding to authors who have been potentially influenced by several authors and have influenced few

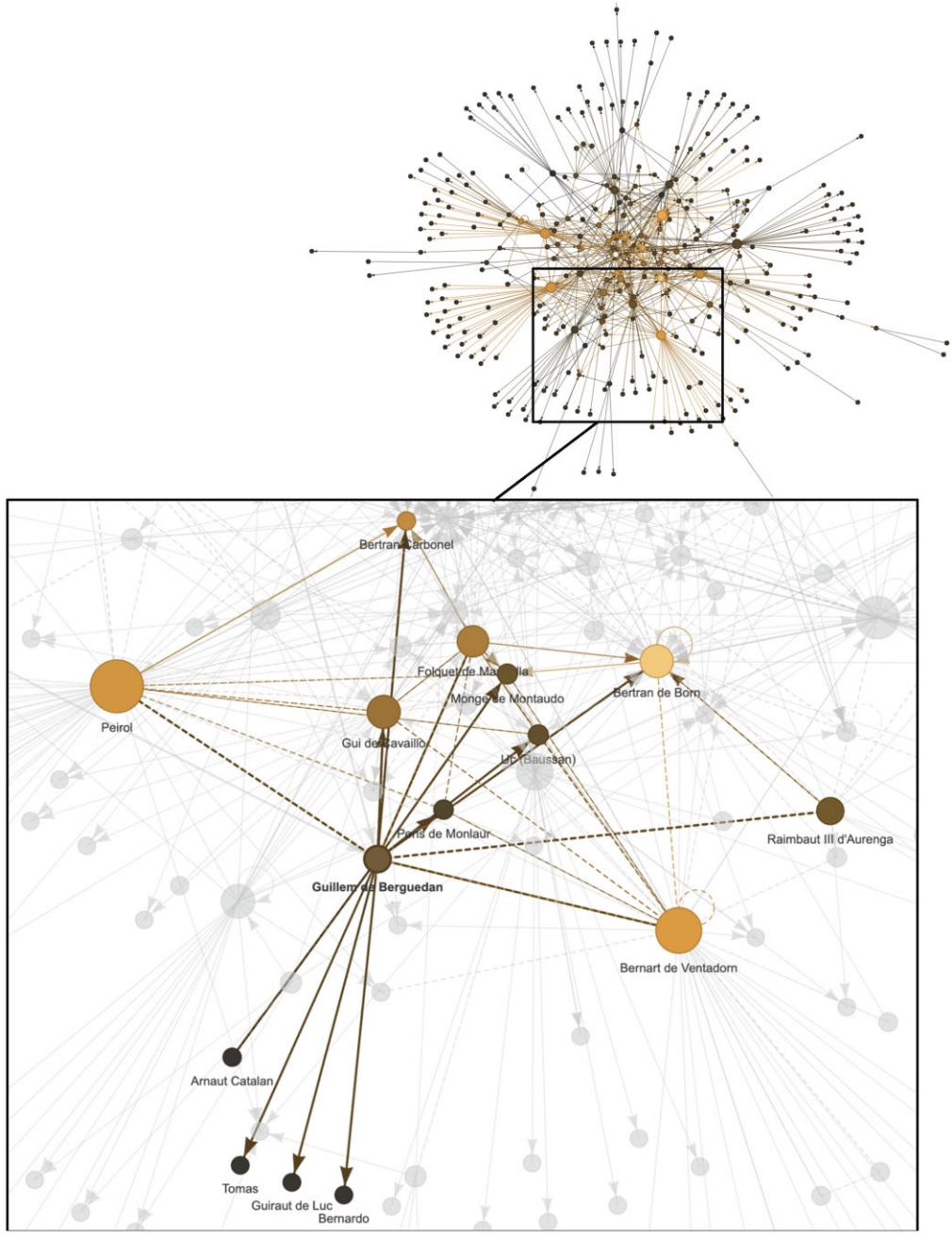

Fig. 3. Network of authors' oeuvres and their connections, based on the practice of contrafactum (solid lines) and on general metrical analogies (dashed lines). The size of a node is proportional to the number of contrafacta the corresponding author has inspired (plus metrical analogies), while its colour represents the number of contrafacta that the author's body of work has produced (plus metrical analogies): a lighter tint denotes larger values while darker shades denote smaller values. In the box, a zoomed-in view of the lower area in which the connections of one of the authors, Guillem de Berguedan, have been highlighted for further clarity

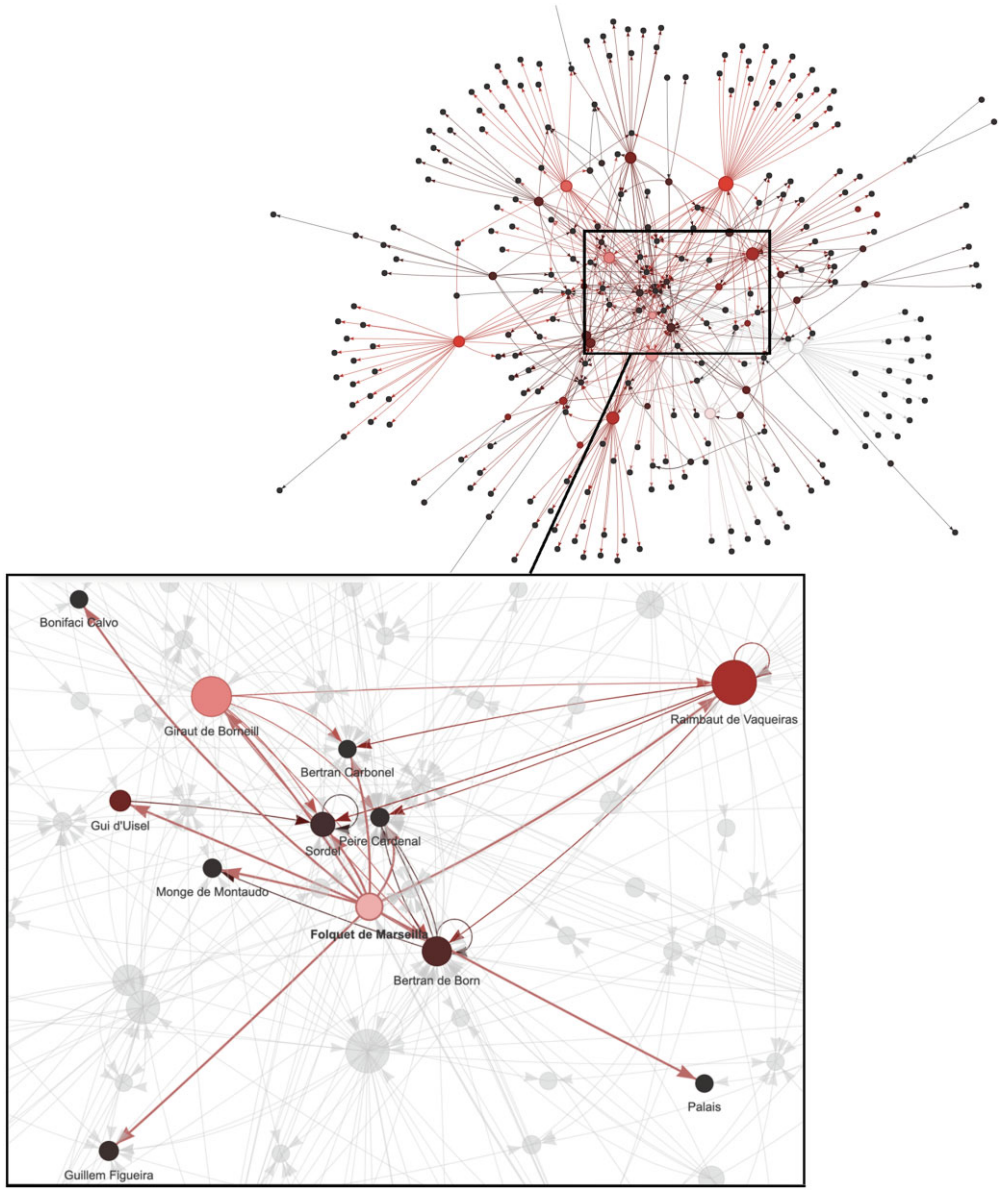

Fig. 4. Network of authors' oeuvres and their connections, based on the practice of contrafactum. The size of a node is proportional to the number of contrafacta the corresponding author has produced, while its colour represents the size of its time-respecting reachability set, that is the number of authors the author may have directly or indirectly inspired with his work. In the box, a zoomed-in view of the central area in which the connections of one of the authors, Folquet de Marseilla, have been highlighted

of them, and we can see that these are mostly late authors. On the bottom right corner, we have authors who, on the contrary, have influenced a great deal of other authors but have not been influenced by many, and here we find mostly early authors. The authors on the top right corner such as Bernart de Ventadorn and Peire Cardenal can be considered as the crucial authors in the transmission of information in the

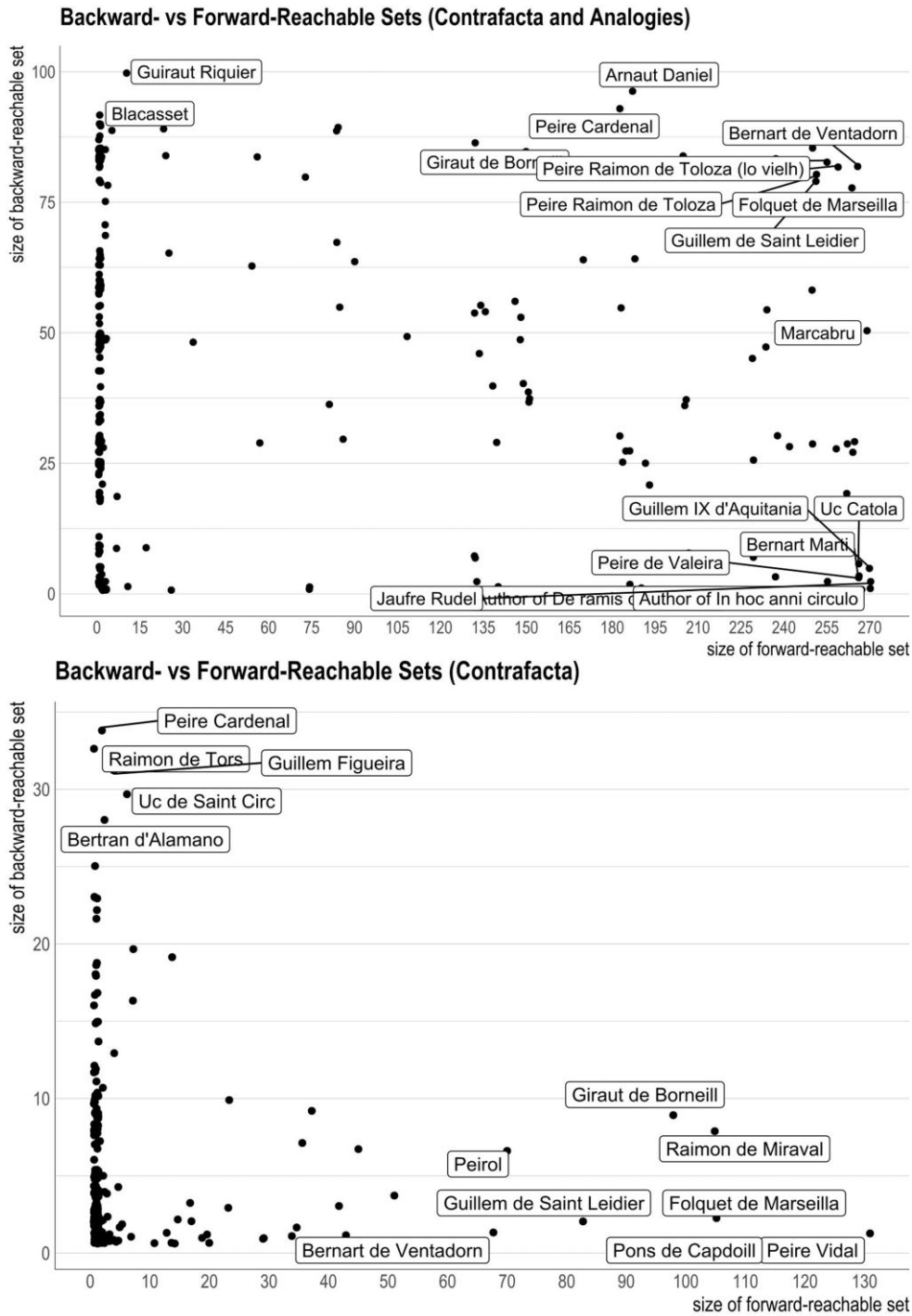

Fig. 5. Size of the backward and forward reachable sets for authors. Each dot corresponds to an author. Its position on the horizontal axis denotes the number of authors to which it is connected via a path-respecting path (i.e. the size of its

dataset. In any case, a great number of authors sit somewhat in the middle of the plot. When we consider contrafacta only, the graph is rather more extreme: most authors have a very small forward-reachable set (i.e. their works have inspired few contrafacta or none at all) and very few have left a greater impact. Some authors, including Peire Cardenal and Raimon de Tors, have a large backward-reachable set, that is to say, they imitated numerous compositions or have imitated songs which in turn were written by prolific authors of contrafacta. This is in line with Peire Cardenal's extensive production of sirventes, and Raimon de Tors' imitation of one of his lyrics. It is also remarkable that in most cases, authors who have large forward-reachable sets have small backward reachable sets and vice-versa, indicating that the most influential authors may have been the most original.

4. Conclusion and Future Work

The scope of this paper is that of presenting our early results on a novel way of visualizing and studying the connections between Provençal troubadours. Our main innovation lies in the temporal network of authors' oeuvres, in which a pair of authors' corpora are connected by a directed link if the latter includes a song which is a contrafactum of a song in the former, and by bidirectional links if one or more of their songs share more generally some metrical similarities.

We have shown how these networks, incorporated into an interactive widget, can be used to assist the exploration of a vast relational dataset of authors and their work, displaying essential information on the authors' compositions, and visualizing at the same time the direct and potential indirect influence of each author in information transmission across the network.

Some of the assumptions generally accepted by medievalists have been confirmed, such as the receptivity of the three most prolific authors of contrafacta, Peire Cardenal, Bertran de Born, and Bertran Carbonel. However, the networks highlight some less obvious features of troubadour tradition. The most relevant names in terms of number of connections are those of Peire Vidal and Peirol: these two poets are comparable in terms of the large diffusion of their songs as they occur in medieval manuscripts, although the former is significantly more productive; their comparable out-degree might appear striking, considering the higher prominence that troubadour studies have accorded to Peire Vidal. Compared with these two authors, celebrated troubadours Bernart de Ventadorn, Gaucelm Faidit, Arnaut de Maroill, Pons de Capdoill, Raimbaut de Vaqueiras, and Raimon de Miraval present smaller, albeit still significant, degrees of influence. Moreover, in comparing authors' reachability and in-degree, we can observe a tendency towards polarization into two main groups of 'propagators' and 'receptors', as the most influential troubadours are shown to be generally less prone to produce contrafacta themselves. Peculiar exceptions are Raimbaut de Vaqueiras, Raimon de Miraval, Giraut de Borneill, and most significantly Bertran de Born, who appears to have been both a prolific imitator and an influential model.

Although all the visualizations here analysed provide a valid perspective for the appreciation of different aspects of the relationship connecting the data, the network representing out-degree and temporal reachability of sole contrafacta (Fig. 4) may be indicated as an optimal balance between readability (by using the simple out-degree measurement), richness of information (by including a chronological dimension), and accuracy (by excluding cases labelled as metrical analogies).

The model presented is a work in progress and we are planning several expansions that will help shed further light on the structure of the dataset and the *Connecting Medieval Music* database in general. The first clear and ongoing extension is the inclusion in

Fig. 5. Continued
reachability set); its position on the vertical axis indicates the number of authors from which time-respecting paths can lead to it, or, in other words, how many authors include it in their own reachability set. On the top, we see the measures computed for the network including both connections representing contrafacta and more general metrical analogies, while on the bottom we have the measures for the network including contrafacta only

our dataset of the works of authors of later troubadour generations presented in the BEdT, as well as authors of lyrics in other languages (French 'trouvères', Galician–Portuguese 'trobadores', and German 'Minnesänger'), as well as connections with the Latin repertoire. Including in the dataset other types of relationship, such as those between authors who have composed a tenso together or whose works are linked by proven intertextual ties, would also be a way to enhance the complexity and detail of our networks. Another ongoing expansion is the analysis of other forms of textual relations in the database, such as tensos, through both a static and temporal framework. Other ongoing directions for expansion include the investigation of the often-vague spatial information about the authors' place of activity, with the aim to better trace their itineraries on the basis of their (time-respecting) interactions with other authors, and the detection of the time-varying communities of authors who share very strong ties among themselves, but are less connected outside of their group, corresponding to 'intellectual communities' which would allow the complex structure of the network to be disentangled (see, e.g. Tantipathananandh *et al.*, 2007; Palla *et al.*, 2007; Pietilänen, and Diot, 2012; Kauffman *et al.*, 2014; He and Chen, 2015; Sekara *et al.*, 2016; Rossetti and Cazabet, 2018). Another direction of expansion is the creation of more sophisticated network models for our database. One very promising option is presented in Vega and Magnani (2018) and Vega and Magnani (2019), who analyse exchange of messages between individuals through a bipartite temporal network in which the links are weighted according to information on the textual similarity between two messages. We believe a similar approach with authors as individuals and songs as messages could prove quite illuminating, provided a sensible measure of similarity between the songs is defined. This would be particularly useful in giving a more detailed account of what we have here gathered in the category of 'metrical analogies'; indeed, our understanding of the system would be improved by specifying (when philological research or mere chronological assumptions allow it) the direction of the influence between two compositions, as well as assigning a degree of certainty to the single connections.

Other directions that will be explored in the future include the problem of network reconstruction (Newman, 2018; Peixoto, 2019), that concerns the possibility to infer the existence of connections (in this case contrafacta or metrical analogies) that are not part of the database, on the basis of networks properties. This is an active field in the case of static network, but there is growing interest towards its generalization to temporal networks (Holme and Saramäki, 2019).

Authors' Contributions

S.M. is responsible for the creation and optimization of the dataset, and the sections pertaining to literary aspects and the interpretation of the networks (primarily in Sections 1, 3.1, 3.2, and 4); M.M. is responsible for writing the code and for the sections pertaining to the construction of the networks (primarily in Sections 2, 3.3, 3.4, and 3.5 and the Supplementary information).

Funding

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 713683 (WIRL); and the Institute of Advanced Study, University of Warwick, award no. 23095-21.

Supplementary Data

Supplementary data are available at DSH online.

References

Antonelli, R. (1984). *Repertorio Metrico della Scuola Poetica Siciliana*. Palermo: Centro di studi filologici e linguistici siciliani.

Asperti, S. (1991). Contrafacta provenzali di modelli francesi. *Messana*, **8**: 5–49.

Asperti, S. and de Nigro, L. (2012). *Bibliografia Elettronica dei Trovatori—v. 2-5-2012*. Available at: http://www.bedt.it/BEdT_04_25/ (accessed 9 February 2021).

Avalle, D. S. and Leonardi, L. (1993). *I manoscritti della Letteratura in Lingua D'oc*. Turin: Einaudi.

Atkin, R. H. (1977). *Combinatorial Connectivities in Social Systems: an Application of Simplicial Complex Structures to the Study of Large Organizations*. Basel: Birkhäuser.

Batagelj, V., Doreian, P., Ferligoj, A., and Kejzar, N. (2014). *Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution*. Chichester: John Wiley & Sons.

Bavelas, A. (1948). A mathematical model for group structures. *Human Organization*, **7**(3): 16–30.

Beauchamp, M. A. (1965). An improved index of centrality. *Behavioral Science*, **10**(2): 161–3.

Bender-deMoll, S. and Morris, M. (2016). *tsna: Tools for Temporal Social Network Analysis*. R package version 0.3.1. Available at: URL https://CRAN.R-project.org/package=tsna.

Billy, D., Canettieri, P., Pulsoni, C., and Rossell, A. (2003). *La lirica galego-portoghese. Saggi di metrica e musica comparata*. Rome: Carocci.

Billy, D. (2002). Hasard et intertextualité: à propos d'un cas de contrafacture (Arnaut de Mareuil, La grans beutaz e.l fis ensenhamens), *Vettori e Percorsi Tematici nel Mediterraneo Romanzo: Convegno; L'Apollonio di Tiro nelle Letterature Euroasiatiche dal Tardo-antico al Medioevo: Seminario, Atti [Roma (Villa Celimontana), 11–14 ottobre 2000]*. Soveria Mannelli: Rubbettino, pp. 89–115.

Boccaletti, S., Bianconi, G., Criado, R. et al. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, **544**(1): 1–122.

Chambers, F. M. (1952). Imitation of form in the old provençal lyric. *Romance Philology*, **6**: 104–20.

Chapman, K. E. (2020). *Network Graph: Models and Imitators in the Troubadour Repertoire*. Available at: http://troubadourmelodies.org/networks/borrowing/network/ (accessed 9 February 2021).

Frank, I. (1952). *Trouvères et Minnesänger*. Saarbrücken: West-Ost-Verlag.

Frank, I. (1953–1967). *Répertoire Métrique de la Poésie des Trobadours*. Vol. **2**. Paris: Champion.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, **40**(1): 35–41.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, **1**(3): 215–39.

Gennrich, F. (1956). *Lateinische Liedkontrafactur*. Darmstat.

Gennrich, F. (1965). *Die Kontrafaktur in Liedschaffen des Mittelalters*. Langen bei Frankfurt.

Gil-Mendieta, J. and Schmidt, S. (1996). The political network in Mexico. *Social Networks*, **18**(4): 355–81.

Guida, S. and Larghi, G. (2014). *Dizionario biografico dei trovatori*. Modena: Mucchi.

He, J. and Chen, D. (2015). A fast algorithm for community detection in temporal network. *Physica A: Statistical Mechanics and its Applications*, **429**: 87–94.

Holme, P. (2005). Network reachability of real-world contact sequences. *Physical Review E*, **71**(4): 046119.

Holme, P and Saramäki, J. (2019). A map of approaches to temporal networks. In Holme, P. and Saramäki, J. (eds), *Temporal Network Theory*. Cham: Springer, pp. 1–24.

Kauffman, J., Kittas, A., Bennett, L., and Tsoka, S. (2014) DyCoNet: a gephi plugin for community detection in dynamic complex networks. *PLoS ONE*, **9**(7): e101357.

Lannutti, M. S. (2008). Intertestualità, imitazione metrica e melodia nella lirica romanza delle origini. *Medioevo Romanzo*, **32**(1): 3–28.

Latora, V. and Marchiori, M. (2007). A measure of centrality based on network efficiency. *New Journal of Physics*, **9**(6): 188.

Marchiori, M. and Latora, V. (2000). Harmony in the small-world. *Physica A: Statistical Mechanics and its Applications*, **285**(3-4): 539–46.

Marshall, J. H. (1980). Pour l'étude des contrafacta dans la poésie des troubadours. *Romania*, **101**: 290–335.

Mazzamurro, M. (2022). *MatteoMazzamurro/Contrafacta-Networks: Contrafacta-Networks-First-Three-Generations (v1.1.0)*. Zenodo. https://doi.org/10.5281/zenodo.6363437.

Medo, M., Cimini, G., and Gualdi, S. (2011). Temporal effects in the growth of networks. *Physical review letters*, **107**(23): 238701.

Milonia, S. (2016). *Rima e Melodia Nell'arte Allusiva dei Trovatori*. Rome: Nuova Cultura.

Mölk, U., and Wolfzettel, F. (1972). *Répertoire Métrique de la poésie lyrique Française des Origines à 1350*. München: Fink.

Newman, M. E. (2018). Estimating network structure from unreliable measurements. *Physical Review E*, **98**(6): 062321.

Nicosia, V., Tang, J., Mascolo, C., Musolesi, M., Russo, G., and Latora, V. (2013). Graph metrics for temporal networks. In Saramäki, J. and Holme, P. (eds), *Temporal Networks*. Berlin, Heidelberg: Springer, pp. 1–26.

Palla, G., Barabási, A., and Vicsek, T. (2007). Quantifying social group evolution. *Nature*, **446**: 664–7.

Peixoto, T. P. (2019). Network reconstruction and community detection from dynamics. *Physical Review Letters*, **123**(12): 128301.

Persico, T. (2017). "Contrefact", "contrafact", "contrafactum" (secoli XIV–XV): falsificazione, imitazione, parodia. *Elephant & Castle*, 5–29.

Pietilänen, A. K. and Diot, C. (2012). Dissemination in opportunistic social networks: the role of temporal communities. In *Proceedings of the Thirteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*. Berlin: ACM, pp. 165–74.

Riolo, C. S., Koopman, J. S., and Chick, S. E. (2001). Methods and measures for the description of epidemiologic contact networks. *Journal of Urban Health*, **78**(3): 446–57.

Rossell, A. (2002). L'intermelodicità come giustificazione delle imitazioni metriche nella lirica trobadorica. In Beggiato F. and Marinetti, S. (eds), *Vettori e Percorsi Tematici nel Mediterraneo Romanzo: Convegno; L'Apollonio di Tiro nelle Letterature Euroasiatiche dal Tardo-antico al Medioevo: Roma, Villa Celimontana, 11–14 ottobre 2000: atti*. Soveria Mannelli: Rubbettino.

Rossetti, G. and Cazabet, R. (2018). Community discovery in dynamic networks: a survey. *ACM Computing Surveys (CSUR)*, **51**(2): 1–37.

Schultze, J. (1989). *Sizilianische Kontrafakturen*. Tübingen: Max Niemeyer.

Schultze, J. (2004). *Amicitia Vocalis: Sechs Kapitel zur Frühen Italienischen Lyrik mit Seitenblicken auf die Malerei*. Tübingen: Max Niemeyer.

Sekara, V., Stopczynski, A., and Lehmann, S. (2016). Fundamental structures of dynamic social networks. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(36): 9977–82.

Solimena, A. (1980). *Repertorio Metrico dello Stil Novo*. Roma: Presso la Società.

Solimena, A. (2000). *Repertorio Metrico dei Poeti Siculo-Toscani*. Palermo: Centro di Studi Filologici e Linguistici Siciliani.

Tang, J., Leontiadis, I., Scellato, S., Nicosia, V., Mascolo, C., Musolesi, M. and Latora, V. (2013). Applications of temporal graph metrics to real-world networks. In Saramäki, J. and Holme, P. (eds), *Temporal Networks*. Berlin, Heidelberg: Springer, pp. 135–59.

Tantipathananandh, C., Berger-Wolf, T., and Kempe, D. (2007). A framework for community identification in dynamic social networks. In: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and DataMining*. New York: ACM, pp. 717–26.

Tavani, G. (1967). *Repertorio Metrico della Lirica Galego-Portoghese*. Rome: Edizioni Dell'ateneo.

Taylor, D., Porter, M. A., and Mucha, P. J. (2019). Supracentrality analysis of temporal networks with directed interlayer coupling. In Holme, P. and Saramäki, J. (eds), *Temporal Network Theory*. Cham: Springer, pp. 325–44.

Vega, D. and Magnani, M. (2018). Foundations of temporal text networks. *Applied Network Science*, **3**(1): 25.

Vega, D. and Magnani, M. (2019). Metrics for temporal text networks. In Holme, P. and Saramäki, J. (eds), *Temporal Network Theory*. Cham: Springer, pp. 147–60.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Vol. **8**. Cambridge: Cambridge University Press.

Whitbeck, J., Dias de Amorim, M., Conan, V., and Guillaume, J. L. (2012). Temporal reachability graphs. In Akan, Ö. B. (ed), *Proceedings of the 18th annual international conference on Mobile computing and networking*. New York, NY: Association for Computing Machinery, pp. 377–388.

Xuan, B. B., Ferreira, A., and Jarry, A. (2003). Computing shortest, fastest, and foremost journeys in dynamic networks. *International Journal of Foundations of Computer Science*, **14**(02): 267–85.

Zhao, D. and Strotmann, A. (2015). Analysis and visualization of citation networks. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, **7**(1): 1–207.

Note

1 Authors of tensos are not directly connected to each other in the network, since, even though their relationship is undisputed, this would imply merging an additional kind of information to those concerning contrafacta and metrical analogies. Creating networks visualizations including information on tensos would however be of interest in the investigation of the relationship between authors, as would the inclusion of many other possible layers of information, such as familiar relationships, acquaintanceship, patronage, or intertextual relationships, quotes, and references.