

Imprint

Université du Luxembourg 2021

Luxembourg Centre for Contemporary and Digital History (C²DH)

Université du Luxembourg
Belval Campus
Maison des Sciences Humaines
II, Porte des Sciences
L-4366 Esch-sur-Alzette

The publication of this special issue was in part supported by the Max Weber Foundation and the Fritz Thyssen Foundation.

Editors

Asst. Prof. Dr. Marten Düring (Luxembourg Centre for Contemporary and Digital History | C²DH)
apl. Prof. Dr. Robert Gramsch-Stehfest (Friedrich-Schiller-Universität Jena)
Dr. Christian Rollinger (Universität Trier)
Dr. Martin Stark (ILS – Institut für Landes- und Stadtentwicklungsforschung, Dortmund)
Clemens Beck, M. A. (Friedrich-Schiller-Universität Jena)

Guest Editors

Dr. Henrike Rudolph, University of Göttingen
Dr. Song Chen, Bucknell University

ISSN 2535-8863

Contact

Principal Contact

JHNR-editors@historicalnetworkresearch.org

Support Contact

Dr. Marten Düring (Université du Luxembourg)
JHNR-support@historicalnetworkresearch.org

Typesetting

text plus form, Dresden, Germany

Cover image

Chinese star chart, British Library, Or.8210/S.3326 recto,
<https://www.bl.uk/collection-items/chinese-star-chart>

Copyediting

Andy Redwood, Barcelona, Spain

Published online at

<https://doi.org/10.25517/jhnr.v5i1>

This work is licensed under a Creative Commons License:
Attribution-NoDerivatives 4.0 (CC BY-ND 4.0)
This does not apply to quoted content from other authors.
To view a copy of this license, please visit
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>



MARCUS BINGENHEIMER

The Historical Social Network of Chinese Buddhism

Journal of Historical Network Research 5 (2021) 233–247

Keywords Chinese Buddhism, Biographies of Eminent Monks, Dharma Drum Person Authority Database

Abstract The Historical Social Network of Chinese Buddhism is a large dataset consisting of c. 17,500 actors and c. 25,000 links. The data was collected between 2007 and 2020 as part of various projects at Dharma Drum, a Buddhist organization in Taiwan. It is based on two main sources: marked-up biographical literature, and the Buddhist Person Authority Database. The main component of the network begins in the late 3rd century and ends with actors in the early 20th.

The network serves as a research tool for various levels of historical reflection. On the micro-level, researchers can explore the ego-networks of persons of interest as they were embedded within the larger Buddhist networks of their time. On the meso-level, researchers can focus on certain periods and discern relevant communities and the communication lines between them. On the macro-level, the network can reveal long-term historical structures and provide quantitative evidence to corroborate or refute previous assumptions about *longue durée* trends within Chinese Buddhism.

This article relates how and why the data was collected, addresses its problems and limitations, as well as its potential for future research. The Historical Network of Chinese Buddhism is extensible, e.g. by including Korean and Japanese actors, or by importing data from the China Biographical Database Project or Wikidata.

The dataset, as well as its sources, is published with documentation on GitHub under a Creative Commons Attribution-Share Alike 4.0 license and is ready for use in open-source social network analysis tools such as Gephi or Cytoscape.

1. Introduction*

In historical research, the word “network” is used both figuratively and in a formal sense.¹ On the one hand, a figurative, informal approach to “networks”, more or less densely theorized, can be useful wherever the focus is on groups, rather than individuals, regions, or periods.² It has the advantage that the network metaphor can be easily expanded to include places and things (like in Actor-network Theory (ANT)), and that data does not need to be in a computable format. In fact, computation does not need to enter the research process at all. The formal approach, on the other hand, requires the systematic collection of computable network data, which can be analyzed with metrics derived from graph theory, and visualized with the help of layout algorithms.³ It produces testable, quantitative evidence to support or refute historical arguments. Arguably, formal network analysis scales better to larger amounts of information over longer time periods and can help to discover patterns that are not revealed by a close reading of sources. As with the other papers in this special issue, the focus here is on formal social network analysis.

* **Acknowledgements:** I am grateful to Laurent Van Cutsem and the editors of this special issue for making valuable suggestions on an earlier draft of this article.

Corresponding author: Marcus Bingenheimer (Temple University), m.bingenheimer@gmail.com

1 For archaeology, the distinction I assert here was articulated by Knappett. Carl Knappett ed., *Network Analysis in Archaeology: New Approaches to Regional Interaction* (New York: Oxford University Press, 2013). Marx mentions the same distinction in his overview of historical network studies. Christian Marx, “Forschungsüberblick zur historischen Netzwerkforschung. Zwischen Analysekategorie und Metapher,” in *Handbuch Historische Netzwerkforschung*, ed. Düring et al. (Münster: Lit Verlag, 2016), 83.

2 This is how the term is used in the history of religions e.g., by Jennifer Eichman for Buddhist networks in the late Ming, or by Anna Collar for cult networks in the Roman Empire. Jennifer Jennifer Eichman, *A Late Sixteenth-Century Chinese Buddhist Fellowship: Spiritual Ambitions, Intellectual Debates, and Epistolary Connection* (Leiden: Brill, 2016); Anna Collar, *Religious Networks in the Roman Empire – The Spread of New Ideas* (Cambridge: Cambridge University Press, 2013).

3 For examples see many of the contributions in Knappett, *Network Analysis*, Tom Brughmans, Anna Collar, and Fiona Coward eds., *The Connected Past – Challenges to Network Studies in Archaeology and History* (New York: Oxford University Press, 2016), and Marten Düring, Ulrich Eumann, Martin Stark, and Linda von Keyserlingk eds., *Handbuch Historische Netzwerkforschung* (Münster: Lit Verlag, 2016).

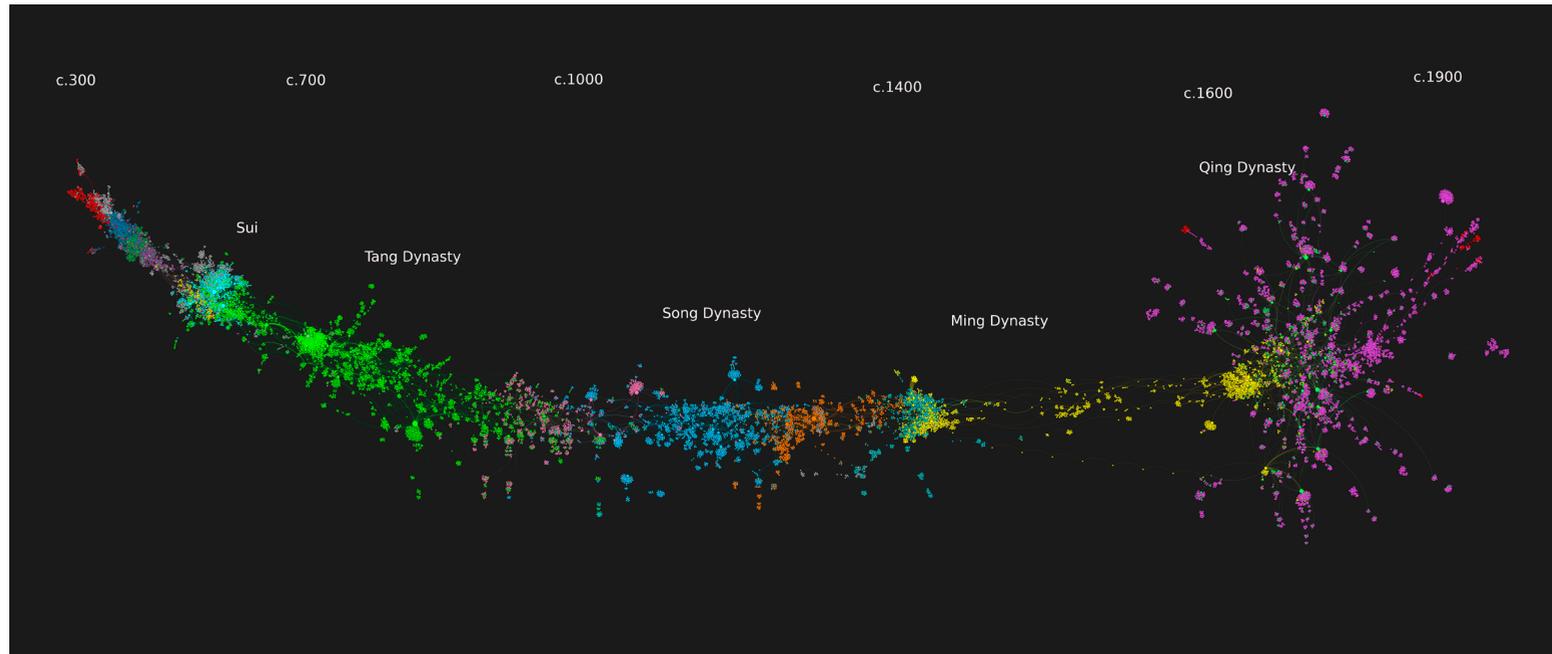


Fig. 1 The main component of the network, laid out and rotated into an approximate timeline (ranging from c. 300–1950 CE).

2. Dataset – Overview and Metrics

The Historical Social Network of Chinese Buddhism is available under a CC Attribution-Share Alike license at https://github.com/mbingenheimer/ChineseBuddhism_SNA. The data was collected between 2007 and 2020 as part of various projects at Dharma Drum, a Buddhist organization in Taiwan. It is based on two main sources: marked-up biographies of eminent Buddhists, and the Dharma Drum Buddhist Studies Person Authority Database (see Sec. 3 below). All links are referenced to a *source text* contained in the CBETA corpus.⁴

As of July 2020, the Historical Social Network of Chinese Buddhism consists of c.17,500 *actors* (nodes) and 25,300 *links* (edges). The latter are the (summed) result of 31,900 recorded connections between single actors (see below for a definition of “connection” in this context).

Edge weights are the sum of the connections between two nodes. They are of little importance in this dataset. Indeed, one single historical interaction can be recorded in various sources (or different passages in the same text) and therefore result in several connections. In other words, a link weight of 5 does not necessarily mean there are records of five *different* interactions between two actors. It might be based on one single event that was recorded in five different passages. Conversely, it might also be the result of two, three, four, or indeed five different interactions. Thus, edge weights have little analytic value in this dataset, beyond serving as a rough indicator for how prominently the interactions between two actors appear in the sources.

The *network density*, usually the fraction between the actual links and all possible links, is also of little use, even for comparative purposes, in diachronic networks that span several generations. By contrast to synchronous networks, in a diachronic historical social network actors can only form connections with contemporaries, which in our case excludes most other actors. Sub-graphs representing different time periods, however, might be gainfully compared with this metric, with their density indicating the level of activity on record for the period.

The *network diameter* of 55 is highly significant as it correlates exactly with the number of generations in the main component. The network diameter is the longest geodesic (i.e., the longest shortest path) through a network. The main component in the Historical Social Network of Chinese Buddhism begins with actors born around 250 CE and runs to actors born around 1950 CE, a period of 1700 years. The main component contains c. 15,500 of the total c. 17,500 actors (~87%), and c. 24,200 of the total 25,300 links (~96%). Dividing the period

4 The CBETA corpus of Buddhist texts is available at <https://github.com/cbeta-git/xml-p5a> (raw data), or at <https://cbetaonline.dila.edu.tw> (user interface).

range of the main component by the network diameter, one arrives at ca. 30 years (1750 years / 55), which is the mean for human biological generations.⁵ That the network diameter reflects the number of generations expected for the time span it attempts to model is encouraging. This indicates that the main component contains no grave errors of the type that connects actors who could not have met, because their life spans did not overlap. A mean of 30 years also models teacher-student generations reasonably well, assuming that teachers were on average between 20 and 40 years older than their students (who were able to receive full-ordination at age twenty).

Diachronic (Historical) spread. The network models interactions of Chinese (and some Indian and Korean) actors associated with Buddhism from the 2nd to the 20th century, i.e., from the Han Dynasty to the People's Republic of China. However, neither the earliest nor the most recent centuries are well represented. No attempt has been made to collect data on post-1911 Buddhism, and most of the early Chinese Buddhists of the late Han and three Kingdoms period (c.150–250 CE) are not connected to the main component. Although the data distinguishes between more than 20 different dynastic periods, including transition periods (e.g. Ming-Qing 明清), almost 80% of the actors are assigned to the following eight periods:

Periods	Approximate percentage of actors in the network
Qing 清 (1644–1911)	26%
Tang 唐 (618–907)	20%
Northern Song 北宋 (960–1127)	10%
Ming 明 (1368–1644)	8%
Southern Song 南宋 (1127–1279)	5%
Five Dynasties 五代十國 (907–979)	4%
Yuan 元 (1279–1368)	3%
Sui 隋 (581–618)	3%

* As defined by the Dharma Drum Person Authority (<http://authority.dila.edu.tw/person/>).

Tab. 1 Percentage of actors by period.

5 Marc Tremblay and H el ene V ezina, "New Estimates of Intergenerational Time Intervals for the Calculation of Age and Origins of Mutations," *American Journal Human Genetics* 66 (2000): 651–658.

Network *centralities* help to identify influential actors. Using an earlier, much smaller version of the network based only on biographies, I have reported how *degree centrality* in biographical literature highlights a different set of actors than *betweenness centrality*.⁶ While the top twenty actors according to degree centrality consisted prominently of translators and patrons, betweenness centrality neatly foregrounded Chan and Vinaya masters. Today, an expanded version of the network contains much more information, especially for the second millennium, and here in particular for the Qing Dynasty. In the current network, the distinction between betweenness and degree centrality is less clear. Betweenness centrality is now somewhat distorted by the relative dearth of information about the mid-Ming. This has inflated the betweenness centrality of some otherwise relatively unknown figures in the 15th and early 16th century who now bridge more densely populated regions of the network before and after the mid-Ming.⁷

The *gender imbalance* in the data is conspicuous. Only about 2% (= c. 390) of the nodes represent women. This imbalance is a reminder that the network is based on sources that, at least with regard to gender, are hardly representative of historical reality (or only in as far as they themselves reflect the inequalities of the past, by what they choose to disregard). There is no reason to assume fewer women than men frequented Buddhist temples and supported Buddhist institutions. Indeed, if the present and the recent past is a guide, female lay-Buddhists seem to be doing more of the day-to-day work in Buddhist communities.

3. Dataset – Collection Principles and History

The Historical Social Network of Chinese Buddhism is extracted via a pipeline of scripts from two sources: marked-up biographies and an authority database data.

3.1 Biographical data extracted from TEI/XML files

Starting in 2007, a team of encoders started to add TEI/XML markup to the biographical literature of Chinese Buddhism. The objective was to mark all person and place names and disambiguate them by mapping them to two evolving Name Authority Databases for person and place names. Also marked were dates, which

6 See Marcus Bingenheimer, “Who was ‘Central’ for Chinese Buddhist History? – A Social Network Approach,” *International Journal of Buddhist Thought and Culture* 28, no. 2 (Dec. 2018): 45–67. The Version 2018-01 contained c. 6,500 actors, while 2020-07 consists of more than 17,500.

7 Research into such neglected figures as Wuji Mingwu 無際明悟 (1381–1446), Bao Yuetan 寶月潭 (d.u.), Zhidu 智度 (1304–1370), and Cuifeng Deshan 翠峯德山 (1468–1549) is much needed.

were standardized via an East Asian Calendar Authority Database.⁸ In addition, the encoders recorded *nexus points*, i.e. TEI markup constructs that encode where people were, when, and with whom. The instructions for the encoders were:⁹

Create a single-person nexus point where the text says:

- Something happened to one single person at a certain time
- Something happened to one single person in a certain place
- Something happened to one single person in a certain place at a certain time

Create a multi-person nexus point where the text says:

- Two (or more) figures interacted (e.g. A becomes student of B, A petitions B, A writes a letter to B, A and B are part of the same group that the emperor presents with purple robes)
- Two (or more) figures interacted in a certain place (or are simply mentioned as being in the same place somehow, which in biographical literature implies they were aware of each other)
- Two (or more) figures interacted at a certain time
- Two (or more) figures interacted in a certain place at a certain time

Table 2 shows which texts were marked-up and how many nexus points were created.

For social network analysis, obviously, only multi-person nexus points are of interest. Only contemporary figures can interact in this sense. Persons and places that are merely mentioned, but not encountered, were not encoded as nexus points (e.g. A remembering person B who visited the same place 200 years earlier, A longing for a place B where he had never been). Though not encoded as a nexus point, mentioned names are still encoded as <persName> or <placeName> in the main body of the text, and might be extracted as part of biographical research into a person.

A special case is the occurrence of non-historical figures in the texts. Some nexus points do include non-historical figures, such as the Bodhisattva Guanyin, who are reported as appearing in dreams or manifesting as human beings. En-

8 For details on the markup see Marcus Bingenheimer, Jen-Jou Hung, and Simon Wiles, "Social Network Visualization from TEI Data," *Literary and Linguistic Computing* 26, no. 3 (2011): 271–278. For the creation of the most comprehensive open database of East Asian Calendar dates see Marcus Bingenheimer, Jen-Jou Hung, Simon Wiles, and Boyong Zhang, "Modeling East Asian Calendars in an Open Source Authority Database," *International Journal of Humanities and Arts Computing* 10, no. 2 (2016): 127–144.

9 Expressed as a pseudo regular expression: (person,[time|place|time,place])|(person,person+,[time|place|time,place]?)

Text	Total number of nexus points	Multi-person nexus points
Collection of Records from the Tripiṭaka 出三藏記集 (卷13–15*) c. 515 CE	474	222
Excerpts from ‘Biographies of Famous Monks’ 名僧傳抄 (1卷) c. 514 CE	188	65
Biographies of Nuns 比丘尼傳 (4卷) c. 511 CE	376	194
(Liang Dynasty) Collection of Biographies of Eminent Monks (梁)高僧傳 (14卷) c. 530 CE	1793	958
Continued Collection of Biographies of Eminent Monks (from the Tang) (唐)續高僧傳 (30卷) c. 665 CE	5478	2179
Song Dynasty Collection of Biographies of Eminent Monks 宋高僧傳 (30卷) c. 988 CE	5204	2244
Ming Dynasty Collection of Biographies of Eminent Monks 明高僧傳 (8卷) c. 1617 CE	592	267
Supplement to the Continued Biographies of Eminent Monks 補續高僧傳 (26卷) c. 1641 CE	3974	1806
New Continued Biographies of Eminent Monks 新續高僧傳 (明清 biographies from 65 卷**) c. 1923 CE	2057	1985
Total	20136	9920***

* Only the chapters containing biographies were marked-up.

** Encoding of parts of this collection is still ongoing. The final tally of nexus points will be slightly higher.

*** The total does not translate directly into network links as nexus points are in effect hyperlinks, which can connect more than two nodes. Due to the limitations of the .gexf file format, these hyperlinks must be represented as individual links. i.e. a single nexus point involving four people must be represented as six edges. (This according to the formula for the number of edges on complete graphs: $n(n-1)/2$, with n being the number of nodes.)

Tab. 2 Nexus points in biographical collections.

counters with non-historical beings, or otherwise legendary connections can distort the network considerably. Such events have been flagged in the XML source and are filtered out during the construction of the dataset.

No attempt is made to provide an ontology of events or interactions, i.e. the nexus points do not record what happened (much less why), merely that something took place (at a certain place, at a certain time). This leaves users at times dissatisfied. In defense, I can only say that creating an ontology of events and implementing it consistently was not a realistic goal during the creation of the dataset. Questions of bandwidth, time, funding, and personal tastes aside, there are basic limitations to modeling historical reality via ontologies. Since the data is openly available, interested researchers are welcome to categorize the c. 20,000 nexus points according to their needs.

3.2 The Dharma Drum Buddhist Studies Person Authority Database

In 2007, the Library and Information Center began to work on creating authority databases to disambiguate dates, person and place names that appeared in various digitization projects at the Dharma Drum library.¹⁰ The Buddhist Studies Person Authority has grown to comprise information on c. 43,000 persons, mainly from China, but also including a sizable number of Japanese, Korean, and Indian figures.¹¹ It was created in tandem with markup projects that needed unique identifiers for people and places, but there were also discrete projects to input reference works such as Yuan Chen 陳垣 (1964), which resolves contradicting birth and death dates in the sources. The Person Authority is by far the largest such database in the field of Buddhist Studies, providing unique IDs, detailed, referenced information about dates, attributed works, and a precis with general information about the person. Especially important for SNA is the recording of lineage. Wherever possible, teachers and students of the person in question are listed. The main source for this information regarding the Ming and Qing is Yūkei Hasebe 長谷部幽蹊 (2008), the best researched, most comprehensive listing of lineage affiliation.¹²

Crucial for assembling the Historical Network of Chinese Buddhism, the data of the Person Authority is available for download under a CC license. This allows researchers to combine the lineage information and other data (e.g. life dates, gender, dynasty) from the Person Authority with the c. 10,000 multi-person nexus points from the c. 3000 TEI/XML files of biographical literature. The networks can be easily merged because they use the same person IDs.

-
- 10 Online at <http://authority.dila.edu.tw/>. Of the three early databases, the East Asian Calendar Database is documented in Marcus Bingenheimer, Jen-Jou Hung, Simon Wiles, and Boyong Zhang, *Modeling*. A fourth, more recent, database allows convenient lookup of catalogs of different canonical editions of the Chinese Buddhist canon. It is a revised version of the Jinglu project (<http://jinglu.beta.org/>) organized by Aming Tu (1953–2016).
- 11 For the current version of the Historical Network discussed here, Japanese and second millennium Korean figures were filtered out. The Person Authority in principle contains a large number of Japanese monks (10,610), connected by lineage data, as well as Joseon Dynasty Korean Buddhists (920), which could be included. The quality of information on non-Chinese actors is, however, not quite as reliable. There are numerous problems with transcription, romanization, use of variant characters in names, etc.
- 12 The tables in this edition should be considered the “final version” of Hasebe’s work, who over the course of 40 years has aggregated information on Ming-Qing Buddhism from a vast range of literature from different genres (“Records of Collected Sayings” (*yulu* 語錄), “Records of Lamp Transmission” (*chuandenglu* 傳燈錄), local gazetteers (*difangzhi* 地方志) etc.).

4. Dataset – Problems and Limitations

Below I will try to assess the limitations of the *dataset*. This is not to be confused with the overall limitations of historical social network analysis as a *method*. Readers of this journal will have their own, informed opinion on the strengths and weaknesses of historical social network analysis.

4.1 Balance

The lineage data from the Person Authority and the biographical information from the markup complement each other with regard to their respective time periods. The five main biographical collections provide relatively dense information until the 10th century, while the lineage discourse provides data for the second millennium.

With the varying organization of the data, however, one would expect to already see a different network structure after the Northern Song (12th century). In fact, the relatively compact band of the diachronic network “explodes” into lineage strands only with the Ming-Qing transition in the 17th century. It remains to be seen exactly why this is the case. In any case, one should remember that information for the two millennia comes from quite different sources, and is structured differently. We have tried to mitigate this to a degree by encoding multi-person nexus points in the biographies of Ming and Qing monks contained in the *Xinxu gaoseng zhuan* 新續高僧傳 (CBETA/B. 151, completed 1923), thus adding some density to the simple tree structure of the lineage data. The addition of the *Xinxu gaoseng zhuan* data was partially successful in that it created greater cohesion for the Qing part of the network. However, even the humongous *Xinxu gaoseng zhuan* could not flesh out the thin network of the mid-Ming (c.1460–1560), the dark age of Chinese Buddhism, about which very little is known. Consequently, our network for this period remains meager.¹³ With the exception of the mid-Ming, the main component of the network is now relatively well balanced between the two millennia.

4.2 Coverage and Completeness (Recall)

Related to *recall* in data mining terms, coverage is the relationship between the amount of SNA relevant information which has been included in the dataset, and the information in principle available to researchers. Total recall would mean

13 For a periodization of the *Xinxu gaoseng zhuan* biographies see Hasebe Yūkei who is among the very few researchers who made use of this work. Hasebe Yūkei 長谷部幽蹊, “Min-Shin jidai kōkai no tenbō – jisatuu no fukkō o megutte” 明清時代教界の展望 – 寺刹の復興をめぐる – [A view on Buddhism during Ming and Qing dynasties – With attention to temple reconstruction]. *Zen Kenkyūjo kiyō* 禪研究所紀要 6 (1976): 189–225.

that by consulting the Historical Network of Chinese Buddhism, one sees all knowable connections that are available to the historian. It is thus a measure of the (in)completeness of a search. For a historical network this is impossible to quantify exactly, but users of the data should be aware of the fact that the incompleteness increases over historic time. The closer we get to the present, the larger the difference is between what was modeled and what could have been modeled as a network from existing sources. This is mainly because there are many more sources for the more recent past.

In its current version, the network models a substantial amount of what is possible to know, in terms of SNA relevant information, about the earlier sections of the main component, from c. 250 CE to c. 500 CE. However, even for this period more links and a few more actors can be found, for instance, in prefaces and colophons to texts, many of which are collected in the catalog section of the *Chu sanzang jiji* 出三藏記集, in the *Essay on Buddhism and Daoism* that is part of the official History of the Wei Dynasty 魏書釋老志, in the *Shishuoxinyu* 世說新語, and in apologetic literature, e.g., as collected in the *Hongmingji* 弘明集. For these centuries, it is in principle possible to survey the existing sources, and perhaps even to arrive at a complete mapping of relationships for this period.

For later times, after say the Sui, there are simply too many sources that record connections between Chinese Buddhists. As with the knowledge of the individual researcher, the network can never hope to be complete. Though this seems obvious and historians always have to work from incomplete, biased sources, users often seem to expect total recall, perhaps conditioned by the use of dictionaries, where one can generally expect to find a word and most of its relevant meanings. Working with the Historical Network of Chinese Buddhism, it is important to be aware that incompleteness of the data increases over time, and there are more additional sources available for the second millennium than for the first.

4.3 Errors (Precision)

Errors in the data can occur on different levels. Firstly, errors introduced at the encoding level include: an encoder might misunderstand a line of classical Chinese and create a nexus point between two actors where there is none; misidentify a person and connect the wrong actors,¹⁴ or they might miss a connection, even though a connection is implied in the text.¹⁵

14 Because of widespread homonymy and the use of abbreviations, this can easily occur. There are e.g. two Xuanzang 玄奘 (Pers. Authority ID: A000294 and A009306) in the Tang dynasty, and the two influential Huiyuan 慧遠 (A001204 and A002000) are only one hundred years apart.

15 This case would simply impact the recall rate (Sec. 3.2), but erroneously, not because data collection was not attempted.

Secondly, the source material might already contain errors. While the encoding accurately represents the meaning of the text, the information is still wrong. This is familiar to historians: just like ourselves, our predecessors sometimes got names and dates mixed up, or they misspell, misidentify, or misunderstand *their* sources. Source errors are problematic, because they are often difficult to detect, and are generally only noticed in the process of active research, not during markup.

Although our dataset probably contains errors of all the above types, there are reasons to be optimistic about the error rate. First, if two persons from different periods are connected erroneously, the network graph will be visibly “bent” or “deformed” by a single connection. In fact, during the course of the project, several encoding errors were found and corrected just by visual inspection of the network. Second, we have checked with the help of programs for “impossible connections” between persons whose life dates do not overlap. This caught both some encoding and some source mistakes.¹⁶ Third, the correlation, mentioned above between network diameter and the expected number of generations for the time period, indicates that the network probably contains few, if any, mistakes that connect two persons of different generations. Fourth, because of the tight authority control, where one single ID is assigned per person, we are confident that each person is indeed represented as only one single node in this network.¹⁷

The main limitation of the current dataset is clearly its recall rather than its precision, i.e., scholars cannot expect the dataset to include all or even most information about links between actors. Especially for later periods, the network offers only a first overview over the social network of an actor, and more focused research will usually be able to uncover more connections.

5. Potential for future use

In spite of its limitations, the dataset is sufficiently large to amount to a completely new tool for the study of Buddhist history in China. It can be directly used in open-source SNA tools such as Gephi or Cytoscape, and the XML of the GEXF version of the data (available on GitHub) can be easily transformed into other graph-data formats.

The most obvious immediate contribution to the research process is that historians can now quickly look up the ego-network of a person of interest, and see

16 Nexus points that reflect source errors are flagged and excluded when the dataset is assembled.

17 Representing the same actor by different nodes in the network, e.g. because of an unrecognized alias, can lead to serious confusion at all levels of inquiry.

who else this person has interacted with. Based on the source references in the edge list, researchers can identify the relevant passages in the sources. In this sense the network functions like a dictionary look-up, querying connections instead of meanings.

The truly innovative power of HSNA, however, lies in its “zoomability.” Far from being restricted to the *micro-level* of ego-networks, researchers can zoom out to the *meso-level* of distinct time periods or cliques of actors, and even further out to contemplate the larger course of Chinese Buddhist history on a *macro-level*, e.g. by comparing network densities between dynasties.

It seems to me that this ability to zoom seamlessly between levels of historical inquiry while remaining within the same methodological framework is unprecedented for the study of Chinese Buddhism. There is no other dataset and methodology that allows the comparison of Buddhist networks in the Tang Dynasty with those of the Ming, and also, within the same tool-chain, the Buddhist network of the Sui emperor Wendi with that of the Liang emperor Wudi. Within the reductionist limits of HSNA as a method, the network perspective is universal in the sense that it can be applied to any period, indeed any tradition, as the papers in the special issue of this journal show.

There are several prospects for the future development of the dataset. One low-hanging fruit for the HSNA informed study of East Asian Buddhism would be to add Japanese and Joseon period (1392–1897) Korean actors. Many of these are already included in the Buddhist Studies Person Authority and can easily be merged in. Both the Korean and Japanese actors form distinct network regions in the second millennium, which are nonetheless connected to the main component. The problem is that the information for the Japanese and the later Korean actors in the Person Authority is not nearly as reliable as the data for Chinese Buddhists, and the network information for them is almost exclusively reliant on lineage charts, leading to relatively simple models. A “Historical Buddhist Network of East Asia”, though of course desirable, should ideally include Korean and Japanese sources beyond the lineage narrative.

Another (relatively) low-hanging fruit is to mine the corpus of Buddhist Temple Gazetteers for HSNA information by adding nexus point markup. This genre might “flesh out” the sparse network of the mid-Ming, as well as pulling the less well connected lineages of the Qing together, so that the Qing part of the network would visually come to more closely resemble the frayed band of the Tang and the Song periods than a blossoming flower.

Beyond simply expanding the Historical Network of Chinese Buddhism lies the challenge of how to merge other datasets into it. While researchers can easily merge smaller networks that result from their own research with the larger network in various ways, larger datasets need to be trimmed down before they can be

usefully included. The two main open source possibilities here are the China Biographical Database (CBDB) and Wikidata.

Connectivity with CBDB has been at least partially achieved by adding CBDB identifiers to the entries of the Buddhist Person Authority. As I have confirmed by testing, the networks can be connected via that crosswalk. In practice, however, only shorter time periods, or ego-networks, can be merged and meaningfully interpreted. Thus both datasets are better reduced first to a certain period or ego-network and then merged.¹⁸

Connectivity to Wikidata has also been achieved, as Wikidata has included the Dharma Drum Person Authority ID as a distinct property.¹⁹ Querying Wikidata with SPARQL can yield interesting network data, which could then be merged with the Historical Network of Chinese Buddhism via the Person Authority ID.

6. References

- Bingenheimer, Marcus, Jen-Jou Hung, Simon Wiles, and Boyong Zhang. "Modeling East Asian Calendars in an Open Source Authority Database." *International Journal of Humanities and Arts Computing* 10, no. 2 (2016): 127–144.
- Bingenheimer, Marcus, Jen-Jou Hung, Simon Wiles, "Social Network Visualization from TEI Data." *Literary and Linguistic Computing* 26, no. 3 (2011): 271–278.
- Bingenheimer, Marcus, Jen-Jou Hung, Simon Wiles, and Boyong Zhang. "Modeling East Asian Calendars in an Open Source Authority Database." *International Journal of Humanities and Arts Computing* 10, no. 2 (2016): 127–144.
- Bingenheimer, Marcus. "Who was 'Central' for Chinese Buddhist History? – A Social Network Approach." *International Journal of Buddhist Thought and Culture* 28, no. 2 (Dec. 2018): 45–67.
- Brughmans, Tom, Anna Collar, and Fiona Coward eds. *The Connected Past – Challenges to Network Studies in Archaeology and History*. New York: Oxford University Press, 2016.
- Chen, Yuan 陳垣. *Shi shi yi nian lu* 釋氏疑年錄 [Record of doubtful dates of Chinese Buddhists]. Beijing: Zhonghua 中華書局, 1964.

18 Luling Huang and I have developed a script that queries an XML representation of CBDB and outputs SNA ready GEXF and GraphML files (<https://github.com/hlulingtu/CBDBto-SNA>). Similar queries could be performed on the SQL data now distributed by the CBDB project.

19 Named "Dharma Drum Buddhist College Person ID" (<https://www.wikidata.org/wiki/Property:P1187>) after the previous name of the college (today called Dharma Drum Institute of Liberal Arts).

- Collar, Anna. *Religious Networks in the Roman Empire – The Spread of New Ideas*. Cambridge: Cambridge University Press, 2013.
- Düring, Marten, Ulrich Eumann, Martin Stark, and Linda von Keyserlingk eds. *Handbuch Historische Netzwerkforschung*. Münster: Lit Verlag, 2016.
- Eichman, Jennifer. *A Late Sixteenth-Century Chinese Buddhist Fellowship: Spiritual Ambitions, Intellectual Debates, and Epistolary Connection*. Leiden, Brill, 2016.
- Grandjean, Martin. “Introduction à la visualisation de données: l’analyse de réseau en histoire.” *Geschichte und Informatik*, no. 18/19 (2015): 109–128.
- Hasebe, Yūkei 長谷部幽蹊. “Min-Shin jidai kyōkai no tenbō – jisatsu no fukkō o megutte” 明清時代教界の展望 – 寺刹の復興をめぐって – [A view on Buddhism during Ming and Qing dynasties – With attention to temple reconstruction]. *Zen Kenkyūjo kiyō* 禪研究所紀要, no. 6 (1976): 189–225.
- Hasebe, Yūkei 長谷部幽蹊. *Min Shin bukkyō kenkyū shiryō (sōden no bu)* 明清佛教研究資料(僧傳之部) [Materials for the Study of Ming and Qing Buddhism (Monastic Biographies)]. Tokyo: Ōbaku bunka kenkyūsho 黄檗文化研究所, 2008.
- Knappett, Carl ed. *Network Analysis in Archaeology: New Approaches to Regional Interaction*. New York: Oxford University Press, 2013.
- Knappett, Carl. “Networks in Archaeology: Between Scientific Method and Humanistic Metaphor.” In *The Connected Past*, edited by Tom Brughmans et al. 21–33. New York: Oxford University Press, 2016.
- Marx, Christian. “Forschungsüberblick zur historischen Netzwerkforschung. Zwischen Analysekategorie und Metapher.” In *Handbuch Historische Netzwerkforschung*, edited by Marten Düring et al. 64–84. Münster: Lit Verlag, 2016.
- Tremblay, Marc, and Hélène Vézina. “New Estimates of Intergenerational Time Intervals for the Calculation of Age and Origins of Mutations.” *American Journal Human Genetics* 66 (2000): 651–658.