Análise de Rede Social e Parentesco no Diyala Paleobabilônico: Pais e Filhos no Arquivo de Nūr-Šamaš

Social Network Analysis and Kinship in the Old Babylonian Diyala. Fathers and Sons in the Archive of Nūr-Šamaš

https://doi.org/10.21814/h2d.3470

Carlos Gonçalves, Universidade de Sπo Paulo, Brasil

Como citar

Gonçalves, C. (2021). Análise de Rede Social e Parentesco no Diyala Paleobabilônico. Pais e Filhos no Arquivo de Nūr-Šamaš. *H2D|Revista De Humanidades Digitais*, *3*(1). https://doi.org/10.21814/h2d.3470

ISSN: 2184-562X

Análise de Rede Social e Parentesco no Diyala Paleobabilônico

: Pais e Filhos no Arquivo de Nūr-Šamaš

Social Network Analysis and Kinship in the Old Babylonian Diyala. Fathers and Sons in the Archive of Nūr-Šamaš

https://doi.org/10.21814/h2d.3470

Carlos Gonçalves, Universidade de São Paulo, Brasil

https://orcid.org/0000-0002-0164-935X

Abstract

The object of the analysis carried out here is the Old Babylonian archive of Nūr-Šamaš, originating from the Diyala region, formed by 116 loan contracts conceded by Nūr-Šamaš, and involving some 400 people as debtors and witnesses. Data obtained from the documents of the archive will be used to generate a graph modelling the relations of the attested persons. The graph will then be split into classes maximizing modularity, that is to say, into clusters optimally composed by personal identifications that statistically tend to connect more inside each cluster than with personal identifications of other clusters. Finally, this will be used to raise or lower the probability of stating that pairs of identifications of the type 'ZZZ, son of YYY' and 'YYY' may correspond respectively to a son and his father.

Keywords

Social Network Analysis; Prosopography; Kinship; Old Babylonian Period; Diyala

Resumo

O objeto da análise levada a cabo aqui é o arquivo paleobabilônico de Nūr-Šamaš, proveniente da região do Diyala, formado por 116 contratos de empréstimo concedidos por Nūr-Šamaš, e envolvendo cerca de 400 pessoas como devedores e testemunhas. Os dados obtidos a partir dos documentos do arquivo serão usados para gerar um grafo modelando as relações entre as pessoas atestadas. O grafo será em seguida particionado em classes que maximizem a modularidade, isto é, em clusters compostos de maneira otimizada pelas identificações pessoais que tendem estatisticamente a se conectar mais no interior de cada cluster do que com identificações pessoas de outros clusters. Por fim, isso será usado para aumentar o diminuir a probabilidade de afirmar que identificações do tipo 'ZZZ, filho de YYY' e 'YYY' posssam corresponder, respectivamente, a um filho e seu pai.

Palavras-chave

Análise de rede social; Prosopografia; Parentesco; Período Paleobabilônico; Diyala

1. Introduction

This paper is the second in a series concerning digital methods to examine sets of documents from the Diyala region in the Old Babylonian period. More precisely, this paper employs research strategies typical to social network analysis and focuses on themes of interest to prosopographical studies. As in the previous paper of the series (Gonçalves, forthcoming), the object of the analysis carried out here is the archive of *Nūr-Šamaš*, which consists of 116 loan contracts concluded by *Nūr-Šamaš* and involves some 400 people as debtors and witnesses.

The archive of *Nūr-Šamaš* was published by Reschid (1965) and Van Dijk (1966). It is composed of the documents numbered 1 to 121 in both publications. Almost all documents are loan contracts in which Nūr-Šamaš is the creditor (the exceptions are Documents 75, 116, 119, 120 and 121). Reschid's transliterations with a few improvements are available online on the Cuneiform Digital Library Initiative's website (henceforth CDLI)[1].

From the point of view of the producers of these documents, the people bound by each transaction were identified by their personal name, with the possible addition of a kinship relation (mostly the father) and a profession. In this paper, the composite made of a personal name and a possible kinship relation and a profession will be called an "identification", and each of its occurrences in specific documents will be called an "attestation".

To the modern scholar, these identifications are not always enough to fully identify the people involved, because the original context is not accessible anymore. One of the consequences of this situation is that it might be difficult or even impossible to assess given identifications, such as "ZZZ son of YYY" and "YYY", and then decide if it would be correct to assume that the two occurrences of "YYY" originally referred to the same person in the documentation.

In what follows, selected data obtained from the documents of the archive of *Nūr-Šamaš* will be used to generate an abstraction, namely a graph of relations. This graph will then be split into clusters maximizing modularity. In other words, as explained below in more detail, a partition of the graph into smaller components will be made so that identifications in the same component will be, as much as possible, more frequently connected among themselves than to identifications from other components. Finally, this will be used to raise or lower the probability of correctly stating that the two occurrences of the name "YYY" such as the ones in the previous paragraph originally referred to the same person or not.

2. The Digital Treatment: Modelling the Data, Building the Graph of Relations and Consolidating Possible Personal Information

Many initiatives with prosopographic aims have applied digital treatment to data from cuneiform tablets. Here is not the place to survey all of them, but a few comments must be made regarding preoccupations of the field that are present in these initiatives. First of all, it is wise to differentiate "evidentiary data", or "data that require only a limited amount of interpretation", from "conclusional data", which "require substantial interpretation", to use the terms employed by the Prosobab database (https://prosobab.leidenuniv.nl/guide.php). In what follows, the data from the archive of *Nūr-Šamaš* will be processed successively by four different computer scripts. The most consequential in terms of producing conclusional data is the third one, while all the others, especially the first two, work in the realm of the evidentiary.

The field is also concerned with whether and how automated computer processes generate conclusional data. Prosobab, for instance, states that it does not "use an automated program for" the purpose of disambiguating attestations, that is to say, merging different attestations of a single individual (https://prosobab.leidenuniv.nl/guide.php). Other initiatives, however, have employed automated programs to solve this issue. An example is the Berkeley Prosopography Services, whose toolkit included "a probabilistic disambiguator that determines the likelihood that two or more name instances refer to the same person" (Schmitz & Pearce, 2013, p. 2). The present research also employs an automated program, in conjunction with the author's considerations, to help identify unique individuals in the documentation.

The very first step of the digital treatment conducted on the archive of *Nūr-Šamaš* was to download the transliterations of the archive's documents from the CDLI website. These transliterations are encoded using only characters of the ASCII character set (in the so-called "ASCII text format", or ATF for short), so they look different from the transliterations in regular Assyriological publications but contain all the pertinent epigraphical information. Document 1 will serve as an example in this paper. The unworked transliteration of Document 1 can be seen in the first column of Table 1. All personal names were then manually marked up with enclosing asterisks. The second column of Table 1 shows Document 1 after this markup.

The transliteration of Document 1 in ASCII text format, before and after the markup of personal names

Transliteration of Document 1 without the markup of personal names	Transliteration of Document 1 w
&P223118 = TIM 03, 001	&P223118 = TIM 03, 001
#atf: lang akk	#atf: lang akk
@tablet	@tablet
@obverse	@obverse
1. __2(disz) gin2 ku3-babbar__	1. __2(disz) gin2 ku3-babbar__
2. __masz2 {d}utu__ u2-s,a-ab	2. __masz2 {d}utu__ u2-s,a-ab
3. __ki__ nu-ur2-{d}utu	3. __ki__ *nu-ur2-{d}utu*
4. __dumu__ ku-bi-ia	4. __dumu__ *ku-bi-ia*
5. {disz}ma-szum	5. {disz}*ma-szum*
6. __dumu__ za-na-tim	6. __dumu__ *za-na-tim*
@reverse	@reverse
1. [a-na] __masz#-gan2__-nim#	1. [a-na] __masz#-gan2__-nim#
2. __ku3#__ u3 __masz2-bi__	2. __ku3#__ u3 __masz2-bi__
3. __i3-la2-e__	3. __i3-la2-e__
4. __igi__ mu-na#-nu-um	4. __igi__ *mu-na#-nu-um*
5. __igi__ sin-sze-me-e#	5. __igi__ *sin-sze-me-e#*
6. __igi__ sin-da-di#-ia#	6. __igi__ *sin-da-di#-ia#*
7. __igi__ ARAD-di#?-ia#	7. __igi__ *ARAD-di#?-ia#*
8. __mu dumu-munus# [eresz]-dingir#__ sza x-x u3 sza {d}iszkur i-ba-ra-a	8. __mu dumu-munus# [eresz]-di
# signs x-x over erasure?	# signs x-x over erasure?

The goal of the next step was to build a table detailing all the identifications of individuals and their roles as creditor (only *Nūr-Šamaš* has this role in the archive), debtors or witnesses in the 116 loan contracts being studied. This was made by feeding the transliterations with marked up personal names into a computer script, which was written by the author in the programing language Python (version 3) and was run inside the Jupyter programming environment. Thanks to the small range of syntactical variations of Old Babylonian loan contracts, the script could be based on only one parsing rule. This rule can be stated in the following, slightly simplified way, where the element inside square brackets is optional and the elements inside curly braces can exchange their order of appearance:

{AMOUNT(S) of COMMODITY(IES) [INTEREST added]; DEBTOR(S); from CREDITOR; took} in front of WITNESS(ES);

or with the Akkadian and Sumerian terms:

{AMOUNT(S) of COMMODITY(IES) [INTEREST uṣab]; DEBTOR(S); ki CREDITOR; šu-ba-an-ti(-meš)} igi WITNESS(ES).

The parsing script also decomposes internally each identification of a DEBTOR, CREDITOR or WITNESS into the elements of the rule PERSONAL NAME [KINSHIP RELATION PERSONAL NAME] [PROFESSION].

Finally, transliterations of personal names were stripped of cuneiform determinatives and editorial marks indicating damaged but identifiable cuneiform signs. Indications of cuneiform signs that were too damaged for reading were not resolved in any way. As an example, the result of the parsing of Document 1 is presented in Table 2.

Identifications attested in Document 1, in ATF

Personal Name	**Kinship**	**Personal Name**	**Role**	**Document Number**	**Face**	**Line**
nu-ur2-utu	son of	ku-bi-ia	Creditor	001	obv.	3
ma-szum	son of	za-na-tim	Debtor	001	obv.	5
mu-na-nu-um			Witness	001	rev.	4
sin-sze-me-e			Witness	001	rev.	5
sin-da-di-ia			Witness	001	rev.	6
ARAD-di-ia			Witness	001	rev.	7

At the same time, a table associating transliterations to normalized personal names was produced. This table was crucial, because, to take an actual example, both ^{d}*nu-ur* ^{d}UTU and $^{\mathrm{d}}$*nu-ur*$_{2}$^{d}UTU are considered writings of the same name: *Nūr-Šamaš*. The table of normalized names and the result of the parser described in the previous paragraphs were fed into a second Python script written by the author, which produced a copy of the set of all the attestations with normalized names instead of the ATF transliterations. The result for Document 1 is shown in Table 3.

Identifications attested in Document 1 with normalized personal names

Personal Name	**Kinship**	**Personal Name**	**Role**	**Document Number**	**Face**	**Line**
Nūr-Šamaš	son of	Kūbiya	Creditor	001	obv.	3
Māšum	son of	Zanātim	Debtor	001	obv.	5
Munānum			Witness	001	rev.	4
Sîn-šeme			Witness	001	rev.	5
Sîn-dādiya			Witness	001	rev.	6
Wardiya			Witness	001	rev.	7

A third computer script was employed to perform a series of tasks with the aim of preparing the data for analysis. The script considered the person-specific fields of each attestation (the identifications) and the number of the documents in which they occur. From the combination of this information, a graph was built: its nodes were the identifications, and two identifications were connected by an edge if, and only if, they occurred together in at least one document. The graph made in this way had 402 nodes and 1879 edges. Figure 1 shows one of its possible visual representations.

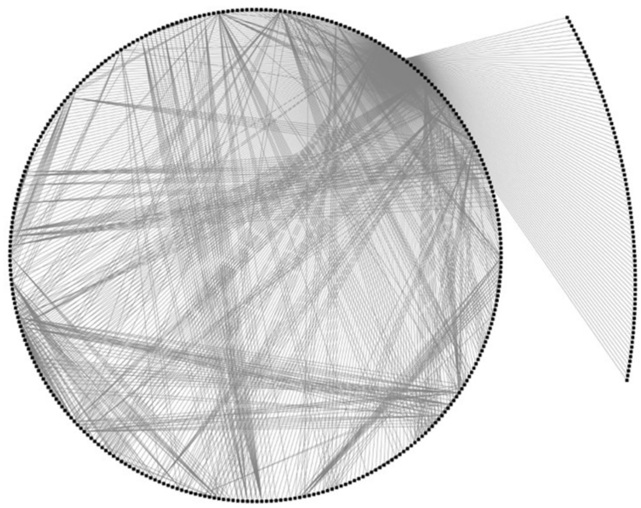

Figure 1: Graph corresponding to all 116 loan contracts of the archive of Nūr-Šamaš. The group of isolated nodes on the right side of the figure correspond to identifications too damaged to be read, so they can only be connected to Nūr-Šamaš

Of course, the graph is an abstraction, and at this step it should not be taken for granted that it preserves social properties of the original community of around 400 people that were involved in the loans registered in the documents. For instance, it is not even assured that each node refers to one person and vice-versa. This particular question was dealt with in Gonçalves (forthcoming).

The next step was based on the fact that graphs can be split into clusters, that is to say, groups of "nodes that are better connected to each other than chance would dictate"[2]. In order to solve this kind of problem, computer scientists developed, among other things, ways of measuring numerically how much a partition of a graph produces clusters. The concept of modularity is one such measure. Explaining modularity in detail involves some mathematics that do not belong here[3]. More than simply partitioning a graph into clusters, the concept of modularity permits one to look for the best way to partition it, that is to say, the one ensuring the graph is split into parts that are clustered as intensely as possible. This is done by seeking a partition that maximizes modularity.

However, the problem of maximizing modularity is a member of a set of excessively time-consuming problems, called NP-complete in computer science jargon (Brandes, 2007, pp. 131-132). Thus, in practice, it is generally only possible to find a partition that produces a sufficiently high modularity. This practical simplification is useful not only for the problem of prosoprographical interest here (disambiguation), but also for a vast range of applications (see again Brandes, 2007). This was done in the present research, where a procedure called "Louvain method", as implemented in the NetworkX package for Python 3, was used to partition the graph[4]. It resulted in a division of the graph into 16 smaller components, the clusters, numbered hereafter 0 to 15. It must be emphasized that each node belongs to exactly one cluster and nodes in each cluster are statistically more connected among themselves than to nodes from other clusters.

Finally, a fourth script was employed to produce an alphabetically ordered list of all identifications and their attestations in the documents of the archive. Three entries of this list are commented on below, to help exemplify how the list is structured. Yet, before proceeding, it must be noted that the usefulness of the list is not based on any visualization of the graph. Rather, it is based on the fact that identifications were distributed optimally into clusters.

Ahūni son of Bazizum.

Ahūni (a-hu-ni) son of Bazizum (ba-zi-zu-um), Doc. 062, obv.4

Belongs to Cluster 4.

Connections in Clusters 4 and 6.

This entry states that the identification “Ahūni son of Bazizum” is attested in document number 62 (obverse, line 4). It also says that, in the splitting of the graph into clusters that maximize modularity, this identification was placed in Cluster 4. Furthermore, the entry indicates that this identification occurs together with identifications that belong to Clusters 4 and 6.

Bēlšunu son of Mērânûm.

Bēlšunu (be-el-szu-nu) son of Mērânûm (me-ra-nu-um), Doc. 008, rev.2

Bēlšunu (be-el-szu-nu) son of Mērânûm (me-ra-nu-um), Doc. 073, rev.7

Bēlšunu (be-el-szu-nu) son of Mērânûm (me-ra-nu-um), Doc. 089, rev.8

Bēlšunu (be-el-szu-nu) son of Mērânûm (me-ra-nu-um), Doc. 090, rev.6

Belongs to Cluster 7.

Connections in Cluster 7.

This entry shows that “Bēlšunu son of Mērânûm” is attested in documents 8, 73, 89 and 90, belongs to Cluster 7 and occurs together with attestations that belong to the same Cluster 7.

Ibnatum son of Apil-Sîn.

Ibnatum (ib-na-tum) son of Apil-Sîn (a-pil-sin), Doc. 053, rev.7

Ibnatum (ib-na-a-tum) son of Apil-Sîn (a-pil-sin), Doc. 079, rev.6

Belongs to Cluster 7.

Connections in Clusters 6 and 7.

Possible father: Apil-Sîn son of Mati-arahtim, in Cluster 10.

Possible father: Apil-Sîn son of Puzuretaya, in Cluster 6.

This entry gathers information about the identification "Ibnatum son of Apil-Sîn", attested in documents 53 and 79. In these documents there are identifications that belong to Clusters 6 and 7. Furthermore, there are two other identifications in other documents that might be useful when trying to identify the father of Ibnatum son of Apil-Sîn. These identifications are "Apil-Sîn son of Mati-arahtim", which belongs to Cluster 10, and "Apil-Sîn son of Puzuretaya", which belongs to Cluster 6.

3. Pairing Fathers and Sons

The rationale behind the idea that clusters maximizing modularity may help improve the likelihood of correctly pairing fathers and sons is that it seems adequate to assume that a son and his father, being members of the same family, may tend to be active in the same circles of acquaintances and, consequently, their identifications may tend to belong to the same cluster or at least have mutual contacts in one same cluster. Reciprocally, attestations of the form "YYY" and "ZZZ father of YYY" that belong to different clusters and that do not even have a contact joining the two clusters to which they belong are less likely to correspond to a son and his father.

So, in the last entry of the list of identifications presented above, the following argument could be made. "Ibnatum son of Apil-Sîn" belongs to Cluster 7 and circulates in Clusters 6 and 7 only. Thus, the likelihood that "Apil-Sîn son of Puzuretaya", which belongs to Cluster 6, identifies his father is bigger than the likelihood of "Apil-Sîn son of Mati-arahtim", which belongs to Cluster 10.

As cases such as this are not rare, the procedure proposed here may be applied in a systematic way. The following entry further exemplifies this strategy.

Ilamliṭṭul son of Ahūni.

Ilamliṭṭul (dingir-lam-li-t,ul) son of Ahūni (a-hu-ni), Doc. 027, obv.4

Ilamliṭṭul (dingir-lam-li-t,ul) son of Ahūni (a-hu-ni), Doc. 031, rev.7

Ilamliṭṭul (dingir-lam-li-t,ul) son of Ahūni (a-hu-ni), Doc. 038, rev.7

Belongs to Cluster 4.

Connections in Clusters 0, 4, and 9.

Possible father: Ahūni son of Bazizum, in Cluster 4.

Possible father: Ahūni son of Warad-x-x, in Cluster 6.

Possible father: Ahūni , in Cluster 9.

The identification “Ilamliṭṭul son of Ahūni “, which is attested in documents 27, 31 and 38, belongs to Cluster 4 and has connections in Clusters 0, 4, and 9. Thus, the chances that “Ahūni son of Warad-x-x”, which belongs to Cluster 6, corresponds to the father of the person identified by “Ilamliṭṭul son of Ahūni” is smaller than those of “Ahūni son of Bazizum”, for it belongs to Cluster 4, and “ Ahūni” of Cluster 9, for it belongs to a cluster where “Ilamliṭṭul son of Ahūni” has connections.

Finally, it is worth considering the following entry, because it exemplifies the limits of the strategy.

Ilīpalûm son of Amtilaba

Ilīpalûm (i3-li2-pa-a-lu-um) son of Amtilaba (am-ti-la-ba), Doc. 044, rev.6

Ilīpalûm (i3-li2-pa-a-lu-um) son of Amtilaba (am-ti-la-ba), Doc. 048, rev.5

Ilīpalûm (i3-li2-pa-a-lu-um) son of Amtilaba (am-ti-la-ba), Doc. 060, rev.8

Ilīpalûm (i3-li2-pa-a-lum) son of Amtilaba (am-ti-la-ba), Doc. 080, rev.6

Belongs to Cluster 2.

Connections in Cluster 2.

Possible father: Amtilaba, in Cluster 6.

In this case, the identification of the possible father “Amtilaba” is part of a cluster to which “Ilīpalûm son of Amtilaba” does not belong (it belongs to Cluster 2) and in which it has no connections (it has connections only in Cluster 2). Hence, the likelihood that “Amtilaba” of Cluster 6 is an identification of the father of the person identified by “Ilīpalûm son of Amtilaba” is not raised by the use of modularity as a measure of clustering. In relation to this specific example, this is confirmed by the fact that “Amtilaba” has indeed connections only in Cluster 6, as its entry in the list of identifications shows:

Amtilaba

Sub-community 6.

Connections in Cluster 6.

Amtilaba (am-ti-la-ba), Doc. 007, obv.5

4. Caveats and Conclusions

It should remain clear that the strategy proposed here is not presented as a definitive and general solution for the problem of pairing fathers and sons. Rather, it is suggested that, especially when large sets of documents are under analysis, the study of clusters maximizing modularity might be a useful piece in the toolbox of the researcher.

Furthermore, this strategy is designed to be used in conjunction with other strategies and heuristics. In particular, it seems to be a wise measure to control the attestations of possible fathers and sons for cases of aliases and homonyms (as presented in Gonçalves, forthcoming).

If the data is appropriately handled and control is kept on the prosopographical hypotheses introduced in the analysis of an archive, it is entirely feasible to design a system in which one would be able to interact with the possibilities of interpretation without getting lost in the wealth of data that even a small archive such as the archive of *Nūr-Šamaš* produces.

Acknowledgements

The author wishes to thank the São Paulo Research Foundation (Fundação de Amparo à Pesquisa do Estado de São Paulo — FAPESP — Grant 2021/01363-6) for funding this research. The author also wishes to thank Robert Middeke-Conlin for the help in correcting and improving the English of this paper. Any remaining flaws are the author's sole responsibility.

Submitted 2021-06-16 | Published 2021-10-31

Notes

1) CDLI. (2020, 21 October) https://cdli.ucla.edu/search/search_results.php?SearchMode=Text&PrimaryPublic However, CDLI states no commitment that search results have stable URLs. The interested reader should fill in the field "Publication" with "TIM 03" in the search form available at home page cdli.ucla.edu.

2) From the leaflet "Network Literacy: Essential Concepts and Core Ideas", published by *Network Science in Education* and downloadable from https://sites.google.com/a/binghamton.edu/netscied/teaching-learning/network-concepts (accessed 14 June 2021).

3) For further thoughts on the issue of community detection and its possibilities, see Newman (2010, pp. 371-373). See also Chen et al. (2015, pp. 240-241) for a comment and Expert et al. (2010) for the original paper on an interesting large-scale application of the concept of modularity.

4) See the article Community detection for NetworkX, available at https://python-louvain.readthedocs.io/en/latest/index.html (accessed 14 June 2021).

References

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikolski, Z., & Wagner, D. (2007). On finding graph clusterings with maximum modularity". In A. Brandstädt, D. Kratsch, & H. Müller (Eds.), Graph-theoretic concepts in computer science. 33rd international workshop, WG 2007. Donburg, Germany, June 2007. Revised papers (lecture notes in computer science 4769) (pp. 121–132). Springer.

Chen, G., Wang, X., & Li, X. (2015). Fundamentals of complex networks. Models, structures and dynamics. Wiley, Higher Education Press.

Expert, P., Evans, T. S., Blondel, V. D., & Lambiotte, R. (2010). Uncovering space-independent communities in spatial networks. Proceedings of the National Academy of Sciences of the United States of America, 108(19), 7663-7668. https://doi.org/10.1073/pnas.1018962108

Gonçalves, C. (forthcoming). Social network analysis: Homonyms and aliases in the old Babylonian Diyala: A study of the archive of Nūr-Šamaš. In C. Gonçalves & C. Michel (Eds.), Interdisciplinary research on bronze age Diyala: Proceedings of the conference held at the Paris Institute for Advanced Study, 25-26 June 2018 (Subartu 47). Brepols.

Newman, M. (2010). Networks. An introduction. Oxford University Press.

Reschid, F. (1965). Archiv des Nūršamaš und andere Darlehensurkunden aus der altbabylonischen Zeit. Inaugural-Dissertation. Heidelberg.

Schmitz, P., & Pearce, L. (2013). Berkeley prosoprography services: Ancient families, modern tools. In F. Tomasi, & F. Vitali (Eds.), Proceedings of the 1st international workshop on collaborative annotations in shared environment: Metadata, vocabularies and techniques in the digital humanities (pp. 1–8). Association for Computing Machinery. https://doi.org/10.1145/2517978.2517980

Van Dijk, J. (1966). The archives of Nūr-Šamaš and other loans. Otto Harrassowitz.