GraphXD

Graphs Across Domains

BIDS GraphXD 2021
March 23, 2021

BERKELEY
Institute for
Data Science

GraphXD
Graphs Across Domains

BIDS GraphXD 2021 • Tuesday, March 23, 2021 • 1:00 – 2:00 PM Pacific

Session 3: NetworkX Application Tutorial

Network Analysis of Ancient Sumerian Texts

Adam Anderson
BIDS Research Program Training Manager

Niek Veldhuis
Professor of Assyriology, UC Berkeley

and the
Sumerian Networks Project Data Science Discovery Team

GraphXD
Graphs Across Domains

BIDS GraphXD 2021 • Tuesday, March 23, 2021 • 1:00 – 2:00 PM Pacific

Session 3: NetworkX Application Tutorial

Network Analysis of Ancient Sumerian Texts

Yashila Bordag
Data Science Discovery Program Student Developer

Colman Bouton
Data Science Discovery Program Student Developer

and the
Sumerian Networks Project Data Science Discovery Team

Sumerian Networks
DS-Discovery Team Presents:

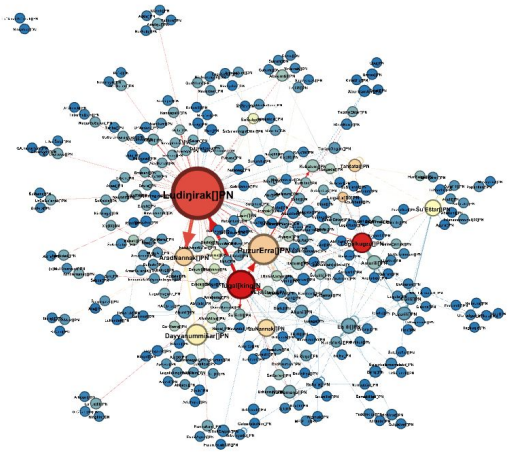
Network Analysis of Ancient Sumerian Texts

The Shoe Archive of Puzrish-Dagan, 2100-2000 BC.

Description: In this demonstration, we will use NetworkX to help solve a riddle contained in a small administrative archive of cuneiform tablets from the ancient Sumerian city-state of Puzrish-Dagan, modern Drehem, Iraq. The archive contains many records of the production of fine shoes, along with precious metals and gems, but why does this small collection of 300 texts exist among thousands of administrative records? To help answer this question, we use network analysis in order to map the relationships between the actors of this small archive, and visualize the social network to find the leaders and their cliques in the archive.

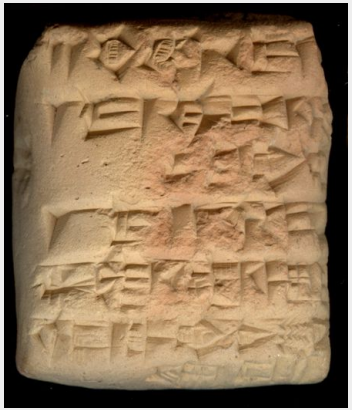
Data Science Discovery Team:

Current Team	Email	GitHub	Role	Semesters
Niek Veldhuis	veldhuis@berkeley.edu	niekveldhuis	Professor, Principal Investigator, developer	9
Adam Anderson	admndrsn@berkeley.edu	admndrsn	Postdoc project manager, developer	9
Yashila Bordag	yashila.bordag@berkeley.edu		Student developer	2
Colman Bouton	clbouton@berkeley.edu		Student developer	1
Dominic Liu	hangxingliu@berkeley.edu	dominicliu129	Student developer	3



The Data in its Various Forms

Cuneiform Tablets



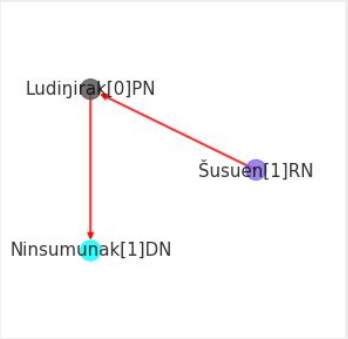
Digitization

a-ru-a
šu-{d}suen
ki
lu₂-dingir-
ra-ta

Lemmatization

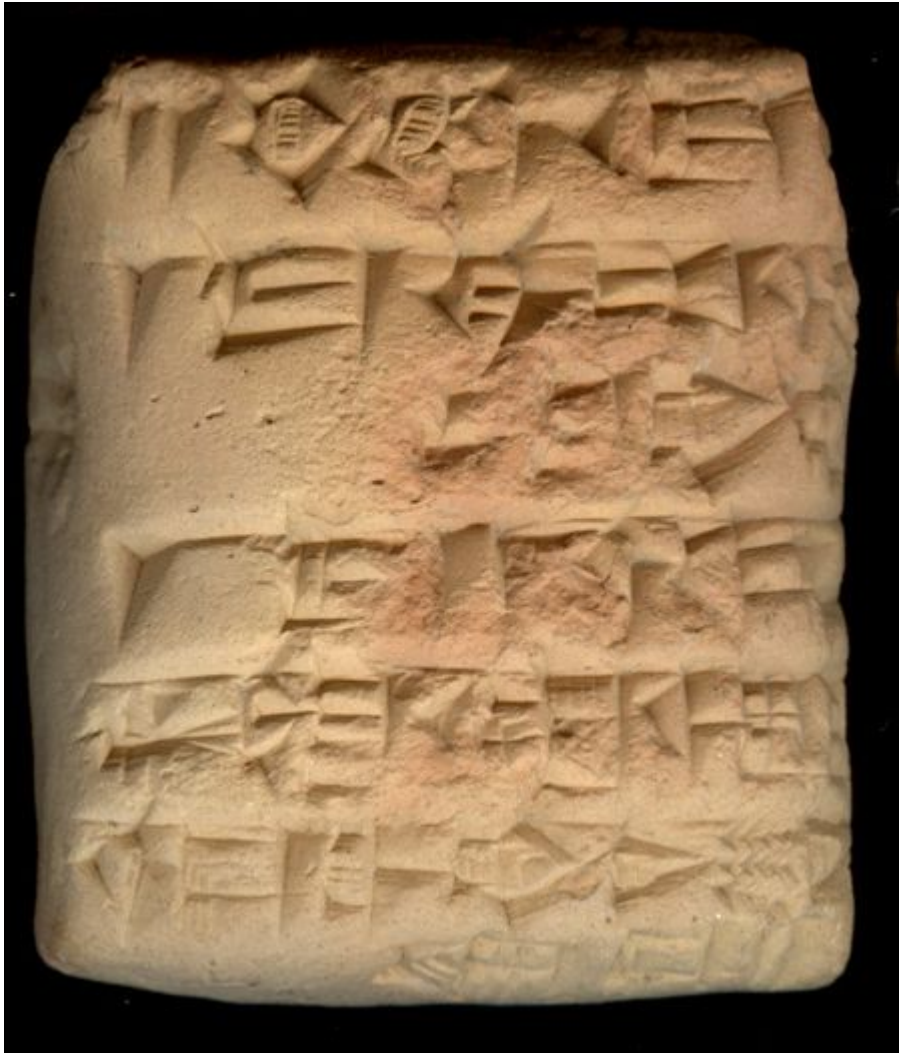
arua[offering]N;
Šusuen[1]RN
ki[place]N;
Ludiŋirak[1]PN

Graph



Tablets

- Puzriš-Dagan: Royal Archive
  - Today: Drehem (south Iraq)
  - Ca. 2060 - 2000 BCE
  - Looted, sold on blackmarket
  - Original arrangement of tablets unknown
  - 15,000 tablets spread over the world
  - Mostly *animals*
  - Treasure & Shoe archive; ca 300 tablets



ORACC: Open Richly Annotated Cuneiform Corpus

- Lemmatization of tablets performed by scholars in the field
- ORACC provides standards for *transliteration* and *lemmatization*.
  - Tablets assigned unique id number
  - Data and metadata can be downloaded in JSON format

```
2 har ku3-babbar 10 gin2-ta
```
n; har[ring]N; kugbabbar[silver]N; n; giŋ[unit]N
2 silver rings 10 shekels each

```
{d}nin-sun2 u3-suh5{ki}
```
Ninsumunak[1]DN; Usuh[1]SN
For Ninsumun of Usuh

```
a-ru-a šu-{d}suen
```
arua[offering]N; Šusuen[1]RN
Offering of Šusuen (the king)

```
ki lu2-dingir-ra-ta
```
ki[place]N; Ludiŋirak[1]PN
expended by Ludiŋira

Rule-based Generator of Nodes and Edges

Nodes:

- Persons (PN)
- Kings (RN)
- Deities (DN)

Roles:

- Recipient
- Offerer
- Source

Keywords:

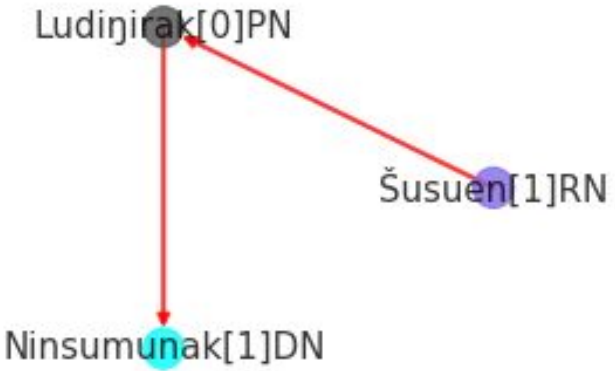
- arua (offering)
- ki (from)

`2 har ku₃-babbar 10 gin₂-ta`
n; har[ring]N; kugbabbar[silver]N; n; giŋ[unit]N
2 silver rings 10 shekels each

`{d}nin-sun₂ u₃-suh₅{ki}`
Ninsumunak[1]DN; Usuh[1]SN
For the goddess Ninsumun of Usuh

`a-ru-a` `šu-{d}suen`
arua[offering]N; Šusuen[1]RN
Offering of Šusuen (the king)

`ki` `lu₂-dingir-ra-ta`
ki[place]N; Ludiŋjirak[1]PN
expended by Ludiŋjira



Resulting directed graph

Data Overview - Nodes

Nodes - people identified as either a source or recipient in the text

- lemma - the word which was identified as an actor in a transaction in the text
- translation, transliteration, pos - metadata from the lemma
- role - role the actor played in a transaction
- attribute - other attributes associated with the node in the text

nodes_frame:

lemma	role	attribute	transliteration	translation	pos
Ludiŋirak[]PN	source		Ludiŋirak		PN
Ninsumunak[]DN	recipient	Usuh[]SN	Ninsumunak		DN

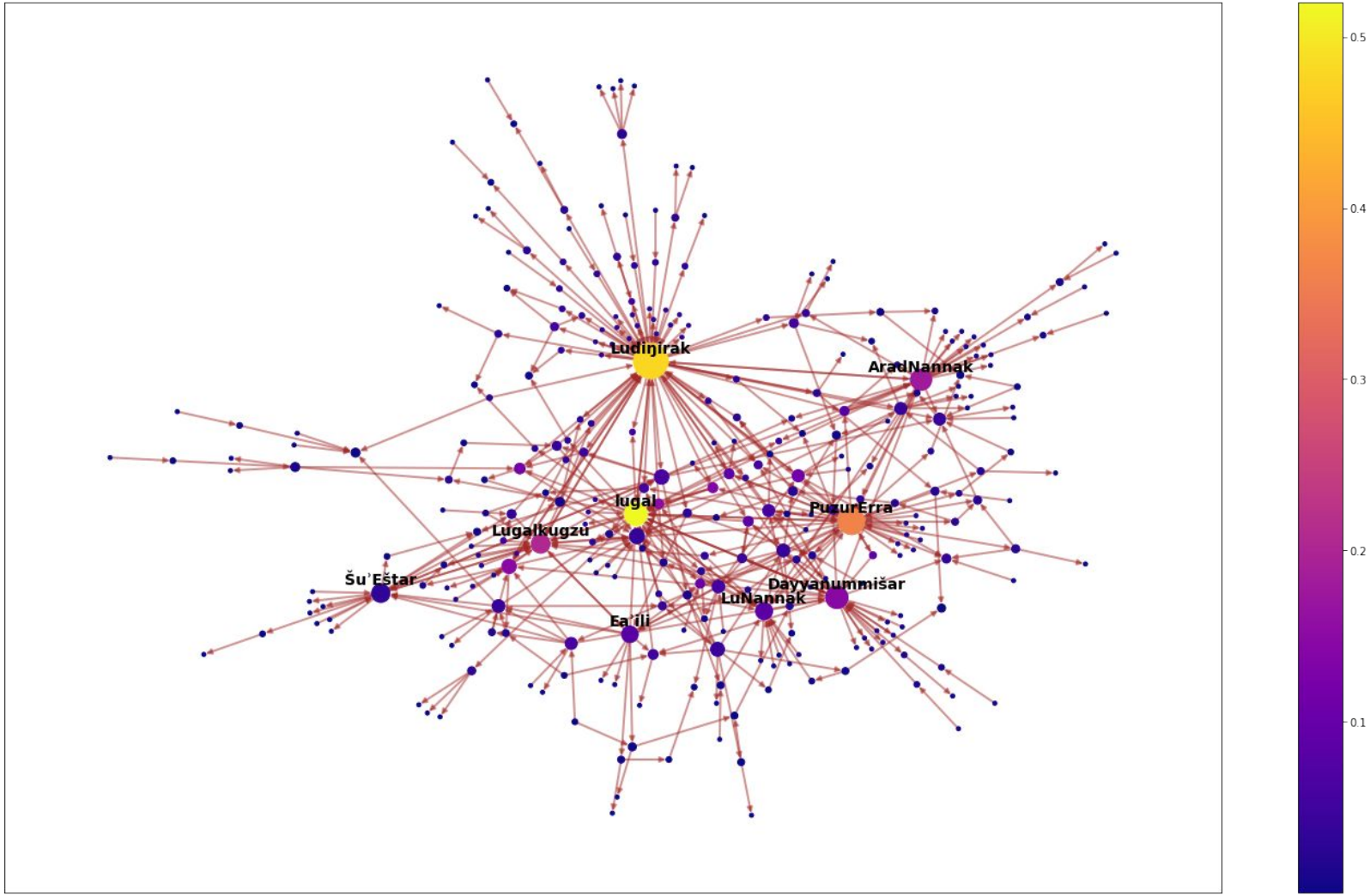
Data Overview - Edges

Edges - transactions which relate different nodes together

- source and target defines directionality, source is the source in the text, target is the recipient in the text
- edge_type: how the nodes are related
- id_text: id of the tablet the edge was scraped from
- weight: number of transactions between two individuals

edges_frame:

	source	target	edge_type	id_text	weight
224	Ludiŋirak[]PN	Ninsumunak[]DN	transaction	P103302	1


Ludiŋirak
AradNannak
lugal
PuzurErra
Lugalkugzu
Šu'Eštar
Dayyanummišar
LuNannak
Ea'ili
eigenvector_undirected
0.5
0.4
0.3
0.2
0.1

Network X Demo

https://bit.ly/3lvd9AD

NetworkX Tutorial.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 6:59 PM

+ Code + Text

NetworkX Sumerian Network Tutorial

Authors: Colman Bouton, Niek Velhuis, Yashila Bordag, Adam Anderson

Introduction

In this demonstration, we will use NetworkX to help explore a small administrative archive of cuneiform tablets from the ancient Sumerian site of Puzrish-Dagan, modern Drehem, Iraq.

This archive contains many records of the production of fine shoes, precious metals, and gems, making it a fascinating subject of study. We are left with several questions such as why does this small collection of 300 texts exist among thousands of administrative records?

To help answer this question, we will use network analysis in order to map the relationships between the actors of this small archive, and visualize the social network to find the leaders and their cliques in the archive.

Dataset

This notebook begins with a prescraped dataset which was constructed as follows. First, the tablets were read and digitized, and then lemmatized in Oracc data base. We then took the ORACC data and used rule based NLP to find the actors in each transaction which we then used to construct our graph. Visually, we can understand this as follows: