Frequent Itemset Mining

Melih Demirel

1953139

Voor dit project is het doel het implementeren van een frequent itemset mining-algoritme in de DBLP-dataset. We willen de maximale frequente itemsets in de DBLP-database berekenen. In dit verslag bespreken we de resultaten van ons algoritme, evenals de benadering die we hebben gevolgd bij de ontwikkeling van onze implementatie en de uitdagingen die we onderweg zijn tegengekomen.

Voor de implementatie van het A-Priori-algoritme beginnen we met het verkrijgen van onze data, omdat deze niet direct beschikbaar is. We maken gebruik van de SAX-parser van Python om auteurs uit publicaties te extraheren en deze in het geheugen op te slaan. Tijdens het inlezen van de data komen we echter een probleem tegen waarbij sommige letters uit andere talen, zoals "ü" ontbreken in onze uiteindelijke lijst van auteurs. Dit wordt veroorzaakt door het encodingtype van de dataset op de DBLP-website, namelijk ISO-8859-1. We lossen dit probleem op met behulp van een eenvoudig script genaamd "recode.py" dat een gezipt XML-bestand met encoding ISO-8859-1 omzet naar UTF-8.

Nadat dit probleem is verholpen, moeten we beslissen hoe we de data opslaan. Het is niet efficiënt om het volledige XML-bestand elke keer opnieuw in te lezen, omdat dit veel onnodige informatie bevat. We hebben twee mogelijke opties: ofwel de hele dataset in het geheugen lezen en hieruit de baskets extraheren, of na het eerste lezen van het XML-bestand deze gegevens wegschrijven naar een tekstbestand, waarbij elke regel een basket vertegenwoordigt en auteurs worden gescheiden door puntkomma's. We kiezen ervoor om beide implementaties te ondersteunen, maar voor ons programma maken we gebruik van het geheugen. Als er echter beperkingen zijn met betrekking tot geheugen, kunnen we eenvoudig de gegevens uit het tekstbestand lezen. Dit gaat uiteraard meer tijd kosten.

Na het succesvol inlezen van de data beginnen we met de implementatie van het A-Priori-algoritme. Een belangrijke keuze hierbij is om niet alle mogelijke combinaties te genereren en te controleren of ze geldige kandidaten zijn, maar om alleen combinaties te genereren binnen een basket. Dit zorgt ervoor dat we alleen kandidaten genereren die al aanwezig zijn in de baskets en dus een support hebben van minimaal 1. Deze aanpak is vooral handig bij grote datasets, omdat het voorkomt dat we geldigheidscontroles moeten uitvoeren voor kandidaten die helemaal niet bestaan.

Om te bepalen of een kandidaat-itemset geldig is, dat wil zeggen dat deze uitsluitend bestaat uit frequente items van grootte k-1, ontwikkelen we een efficiënte methode. We schrijven een functie die zorgt voor een vroege uitgang als bijvoorbeeld de eerste deelverzameling al geen frequente itemset is. Hierdoor hoeven we de rest van de deelverzamelingen niet te controleren, wat aanzienlijke tijdswinst oplevert. Deze optimalisatie is vooral handig wanneer er weinig frequente items zijn.

De maximale frequentie van een groep met grootte k wordt eenvoudig bepaald door de maximale frequentie van de kandidaatsets in stap k te selecteren.

Na de implementatie van deze functies moeten we een keuze maken voor de threshold. Als we alle groottes van groepen willen berekenen met de maximale frequentie, kunnen we in principe een threshold van 1 gebruiken. Echter, naarmate de dataset groter wordt, neemt de tijd die nodig is om kandidaatsets in de volgende stap te genereren exponentieel toe, omdat frequente itemsets die we doorgeven aanzienlijk groter kunnen zijn, met inbegrip van alle auteurs.

Daarom besluiten we om een dynamische threshold te gebruiken in plaats van een statische threshold. Deze threshold verandert op basis van de huidige stap, zodat we een lagere threshold kunnen toepassen voor de eerste stap en een andere threshold voor elke volgende stap. De waarden voor deze thresholds, aangeduid als x en y, worden bepaald op basis van het door de gebruiker opgegeven percentage in de dynamicThreshold-functie. Hoe hoger het percentage (gaat meer canditates als frequent toelaten), hoe lager de threshold zal zijn.

Ons programma kan op drie manieren worden uitgevoerd. Het kan 2, 1 of geen argumenten accepteren. Als u een XML-bestand wilt converteren naar een tekstbestand, geeft u de XML-bestandsnaam en de naam van het tekstbestand op als parameters. Als u echter twee parameters opgeeft, voert het programma alleen de conversie uit zonder het A-Priori-algoritme uit te voeren. Na deze eenmalige conversie is het handiger om de gegevens uit het tekstbestand te lezen, omdat dit proces tot wel 5 keer sneller is dan het lezen van het volledige XML-bestand. Bijvoorbeeld, het lezen van het volledige XML-bestand kostte 150 seconden, terwijl het lezen van gegevens uit een tekstbestand slechts 30 seconden duurde. Als derde optie kan het programma een tekstbestand lezen met een standaardbestandsnaam die in het programma is ingesteld. Het in- of uitschakelen van de dynamische threshold kan worden beheerd op regel 191 van het programma.

Dit zijn de resultaten van onze toepassing voor zowel statische als dynamische thresholds met verschillende waarden op DBLP50000.xml dataset. Deze heeft 49232 publicaties en 69697 verschillende auteurs.

Als eerste gaan we onze xml omzetten naar een txt file met de volgende commando:

Python3 main.py dblp50000.xml dblp50000.txt

Vanaf nu gaan we gebruik maken van dblp50000.txt voor ons programma.

Statisch: getest tot threshold 3. Dit heeft 20 minuten geduurd.

```
Static Threshold enabled: 10

-------------------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
-------------------------------------------------------------------------------------
             1 | Not Used       |                  35 | Philip S. Yu, Micha Sharir
             2 |             10 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')

This took: 32.83431935310364 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```

```
Static Threshold enabled: 5

-----------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
-----------------------------------------------------------------------------
             1 | Not Used       |                  35 | Philip S. Yu, Micha Sharir
             2 |              5 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')
             3 |              5 |                   8 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir')
             4 |              5 |                   6 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir', 'Prithviraj Banerjee')

This took: 198.94976902008057 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```

```
Static Threshold enabled: 4

-----------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
-----------------------------------------------------------------------------
             1 | Not Used       |                  35 | Philip S. Yu, Micha Sharir
             2 |              4 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')
             3 |              4 |                   8 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir')
             4 |              4 |                   6 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir', 'Prithviraj Banerjee')
             5 |              4 |                   4 | ('David Botstein', 'J. Michael Cherry', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng')

This took: 347.90817880630493 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```

```
Static Threshold enabled: 3

-----------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
-----------------------------------------------------------------------------
             1 | Not Used       |                  35 | Philip S. Yu, Micha Sharir
             2 |              3 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')
             3 |              3 |                   8 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir')
             4 |              3 |                   6 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir', 'Prithviraj Banerjee')
             5 |              3 |                   4 | ('David Botstein', 'J. Michael Cherry', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng')
             6 |              3 |                   3 | ('Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry
', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Shuai Weng'), ('Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'Kara Dolinski
', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Shuai Weng'), ('Gail Binkley', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'Kara Dolinski',
 'Laurie Issel-Tarver', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Shuai Weng'), ('Gail Binkley', 'J. Michael Cherry', 'Laurie Iss
el-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Laurie Issel-Tarver', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'Lauri
e Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Mark Schroeder', 'Shuai Weng'), ('Gail Binkley', 'J. Michael Cherry', 'Kara
Dolinski', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'Kara Dolinsk
i', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'J. Michael Cherry', 'Kara Dolinski
', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Shuai Weng'), ('J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', '
Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'J. Michael Cherry', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai W
eng'), ('David Botstein', 'J. Michael Cherry', 'Kara Dolinski', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder'), ('Gail Binkley', 'J. Mi
chael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Selina S. Dwight'), ('David Botstein', 'Gail Binkley', 'Kara Dol
inski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolin
ski', 'Mark Schroeder', 'Selina S. Dwight'), ('David Botstein', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight'), ('Catherine A. Ball', 'David Botstein', 'J. Michael Cherry', 'Mark Schroeder', 'Selina S. D
wight', 'Shuai Weng')
             7 |              3 |                   3 | ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Shuai Weng'), ('Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-T
arver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'Kara Dolinski', 'Laurie Iss
el-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'K
ara Dolinski', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng'), ('David Botstein', 'Gail Binkley', 'J. Michael Cherry'
, 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight')
             8 |              3 |                   3 | ('David Botstein', 'Gail Binkley', 'J. Michael Cherry', 'Kara Dolinski', 'Laurie Issel-Tarver', 'Mark Schroeder', 'Selina S. Dwight', 'Shuai Weng')

This took: 1233.7444489002228 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```

Dynamisch:

```
Dynamic Threshold enabled:
0.5% of singles will go to L2.
20% of k size itemsets will go to L(k+1).

------------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
------------------------------------------------------------------------------
             1 |             11 |                  35 | Philip S. Yu, Micha Sharir
             2 |              2 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')
             3 |              2 |                   8 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir')
             4 |              2 |                   6 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir', 'Prithviraj Banerjee')

This took: 31.567779302597046 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```

Opmerking: hetzelfde resultaat is te zien wanneer we gebruik maken van een statische threshold 5 (boven). Dynamisch wordt het 6x keer sneller gevonden binnen deze parameters.

```
Dynamic Threshold enabled:
0.5% of singles will go to L2.
30% of k size itemsets will go to L(k+1).

------------------------------------------------------------------------------
Group Size (k) | Used Threshold | Max Frequency Count | Example Author Groups
------------------------------------------------------------------------------
             1 |             11 |                  35 | Philip S. Yu, Micha Sharir
             2 |              1 |                  20 | ('Irith Pomeranz', 'Sudhakar M. Reddy')
             3 |              1 |                   8 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir')
             4 |              1 |                   6 | ('Alok N. Choudhary', 'J. Ramanujam', 'Mahmut T. Kandemir', 'Prithviraj Banerjee')
             5 |              1 |                   1 | ('Mark H. Overmars', 'Mark de Berg', 'Micha Sharir', 'Pankaj K. Agarwal', 'Sariel Har-Peled'), ('A. Rust', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle'), ('A. Rust', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle'), ('A. Rust', 'J. Smith', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'J. Smith', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle'), ('A. Rust', 'E. Eyras', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'J. Smith', 'M. Hammond', 'S. Searle', 'Y. Chen'), ('A. Rust', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'S. Searle', 'Y. Chen'), ('A. Rust', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'S. Searle', 'T. Down'), ('A. Rust', 'J. Smith', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'S. Searle', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'S. Searle', 'T. Down'), ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'T. Down'), ('A. Rust', 'J. Smith', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'Y. Chen'), ('A. Rust', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'J. Smith', 'M. Hammond', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'T. Down'), ('A. Rust', 'E. Eyras', 'J. Smith', 'T. Down', 'Y. Chen'), ('J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle'), ('E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down'), ('J. Smith', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('E. Eyras', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle', 'T. Down'), ('J. Smith', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'Y. Chen'), ('E. Eyras', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down'), ('E. Eyras', 'J. Smith', 'S. Searle', 'T. Down', 'Y. Chen'), ('J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'Y. Chen'), ('E. Eyras', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down'), ('E. Eyras', 'J. Smith', 'R. Pettett', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'T. Down', 'Y. Chen')
             6 |              1 |                   1 | ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'J. Smith', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle', 'T. Down'), ('A. Rust', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down'), ('A. Rust', 'J. Smith', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'R. Pettett', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen')
             7 |              1 |                   1 | ('A. Rust', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'S. Searle', 'T. Down', 'Y. Chen'), ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen'), ('E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'T. Down', 'Y. Chen')
             8 |              1 |                   1 | ('A. Rust', 'E. Eyras', 'J. Smith', 'M. Hammond', 'R. Pettett', 'S. Searle', 'T. Down', 'Y. Chen')

This took: 707.0575511455536 seconds.
PS C:\Users\Melih\Desktop\BDA TAAK 1>
```