

# **EER→MLN: EER Approach For Modeling, Mapping, And Analyzing Complex Data Using Multilayer Networks (MLNs)**

**Kanthi Komar<sup>1</sup>, Abhishek Santra<sup>2</sup>,  
Sanjukta Bhowmick<sup>3</sup> and Sharma Chakravarthy<sup>4</sup>**

**<sup>1,2,4</sup>Information Technology Laboratory, CSE Department,  
University of Texas at Arlington, Arlington, Texas, USA**

**<sup>3</sup>CSE Department, University of North Texas, Denton, Texas, USA**

**Email: <sup>1</sup>kanthisannappa.komar@mavs.uta.edu, <sup>2</sup>abhishek.santra@mavs.uta.edu,  
<sup>3</sup>sanjukta.bhowmick@unt.edu, <sup>4</sup>sharmac@cse.uta.edu**

---

# Complex Data Analysis: Application Categories



Movies



Actors



Ratings



Genres



Author Collaborations



Publications



Conferences



Years



Highly rated actor groups working in similar genres but have not co-acted together in any movie?

Same Entities

Multiple Relationships

For the most popular collaborators in each conference, the most active 3-year period(s)?

Different Entities

Multiple Relationships



Flight Routes



Author Collaborations



Author Residence



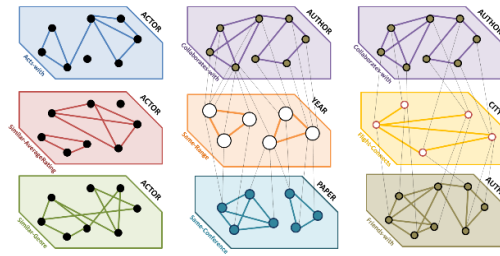
Author Friendship

Best city to hold conferences of authors to maximize attendance?

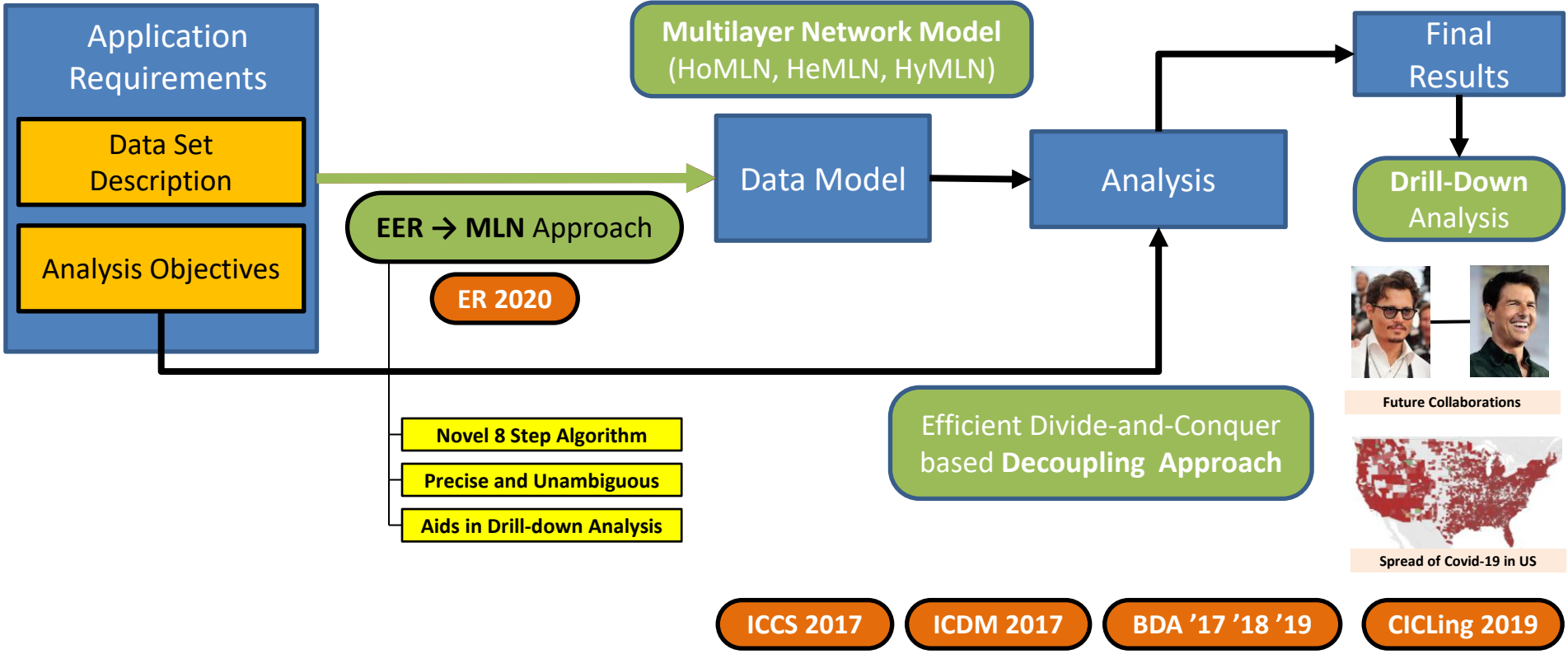
Same & Different Entities

Multiple Relationships

# Big Complex Data Analytics Flow Chart



Multilayer Network Model (HoMLN, HeMLN, HyMLN)

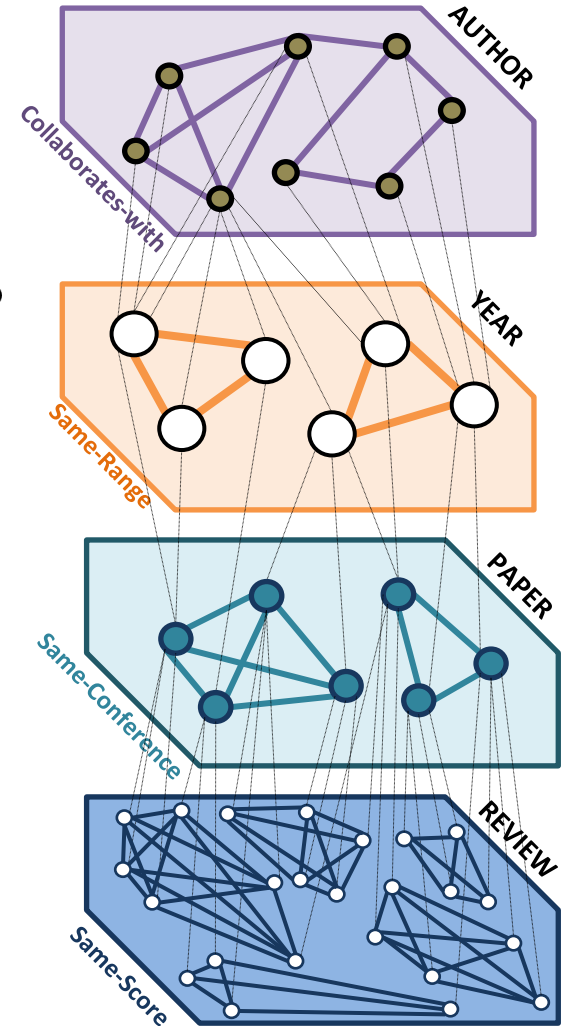
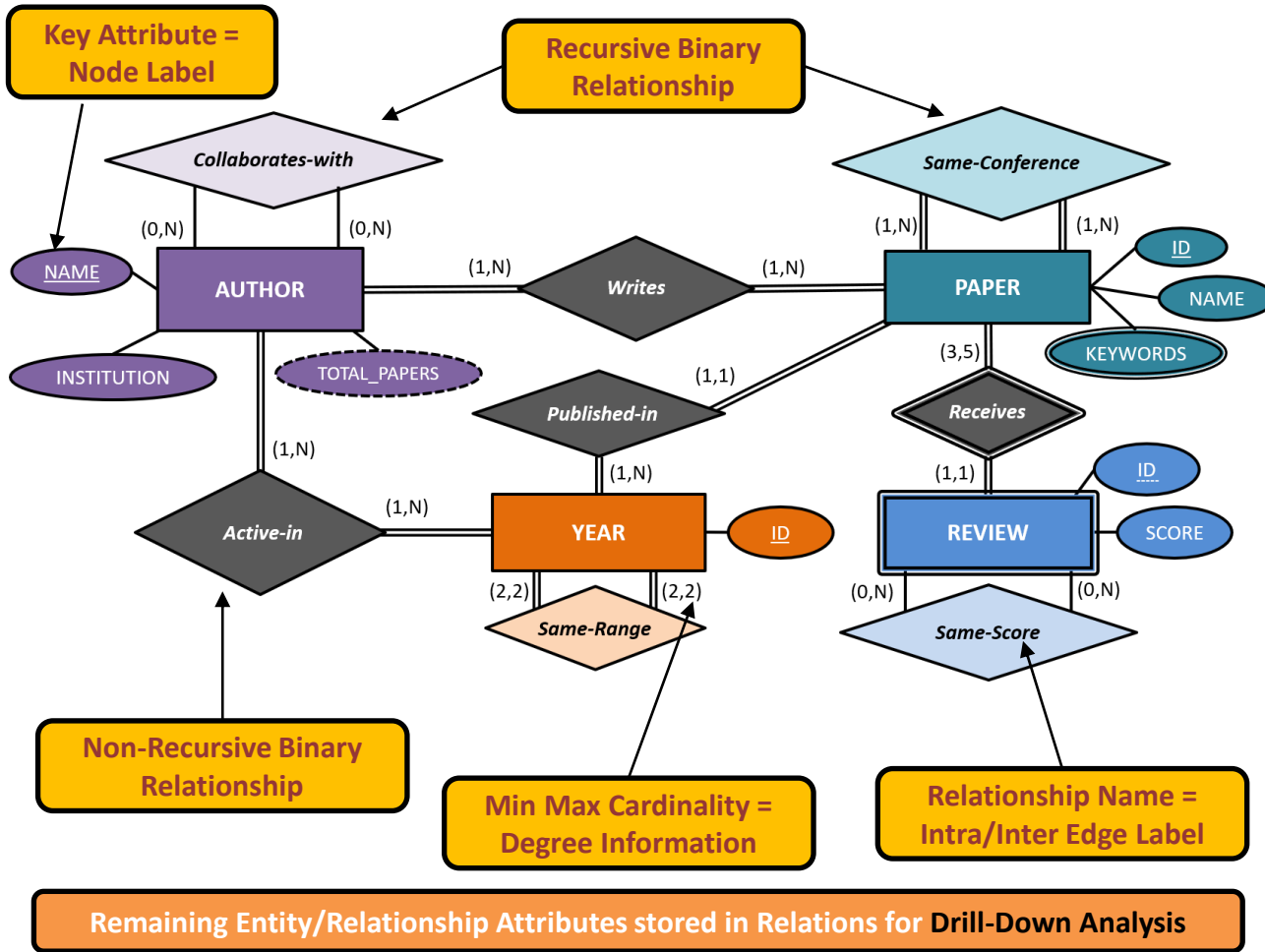


# Data Model: Multilayer Networks (Overview)

- A multilayer network  $\text{MLN}(\mathbf{G}, \mathbf{X})$  is,
    - $\mathbf{G}$  = Set of Simple Graphs
      - $G_i(V_i, E_i)$  represents  $i^{\text{th}}$  layer
    - $\mathbf{X}$  = Set of Bipartite Graphs between layers
      - $X_{i,j}(V_i, V_j, L_{i,j})$ : for  $G_i, G_j$ ;  $L_{i,j}$ : Set of Inter-layer Edges
  - **Homogeneous MLN (HoMLN)**
    - Modeling interactions among same set of entities
    - $V_i = V_j$ , **Implicit** inter-layer edges
  - **Heterogeneous MLN (HeMLN)**
    - Modeling interactions among different sets of entities
    - $V_i \neq V_j$ , **Explicit** inter-layer edges
-

# EER Model → MLN Model: The 8 Step Algorithm

## Research Paper Publication Data Set (DBLP) Modeling

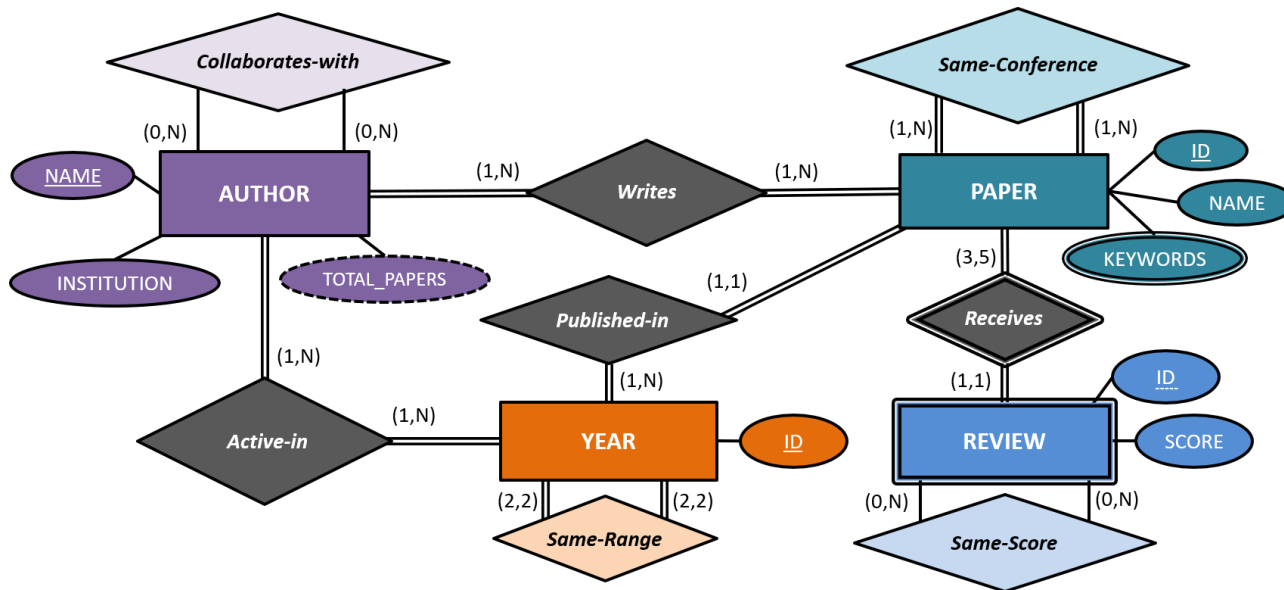


Heterogeneous MLN (HeMLN)

# EER Model → MLN Model: The 8 Step Algorithm

## Research Paper Publication Data Set (DBLP) Modeling

Relations obtained as by product used for Drill-Down Analysis



### Author

<u>Name</u>	Institution
-------------	-------------

### Collaborates-with

<u>Author1Name</u>	<u>Author1Name</u>
--------------------	--------------------

### Paper

<u>ID</u>	Name	PublishYearID
-----------	------	---------------

### Same-Conference

<u>Paper1ID</u>	<u>Paper2ID</u>
-----------------	-----------------

### Keywords

<u>PaperID</u>	Keyword
----------------	---------

### Review

<u>ID</u>	<u>ReviewPaper</u>	Score
-----------	--------------------	-------

### Same-Score

<u>Review1ID</u>	<u>Review2ID</u>
------------------	------------------

### Year

<u>ID</u>
-----------

### Same-Range

<u>Year1ID</u>	<u>Year2ID</u>
----------------	----------------

### Active-in

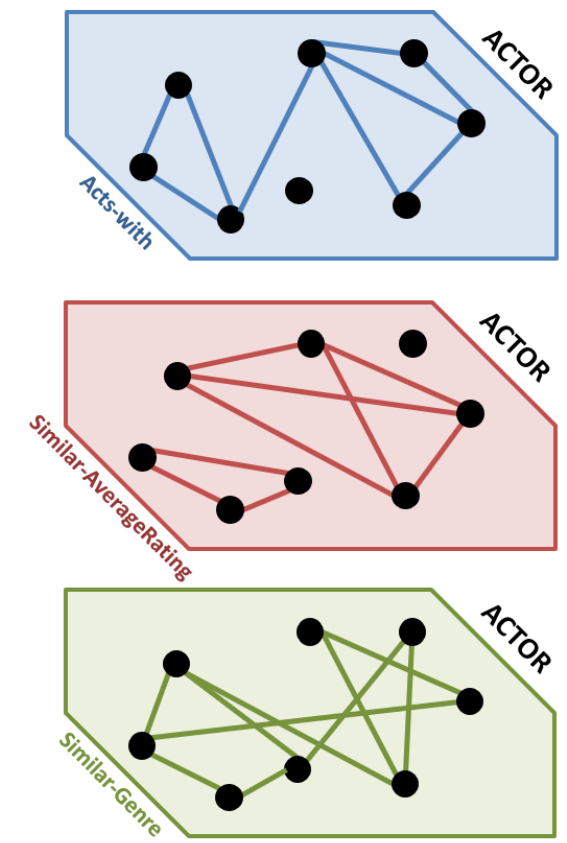
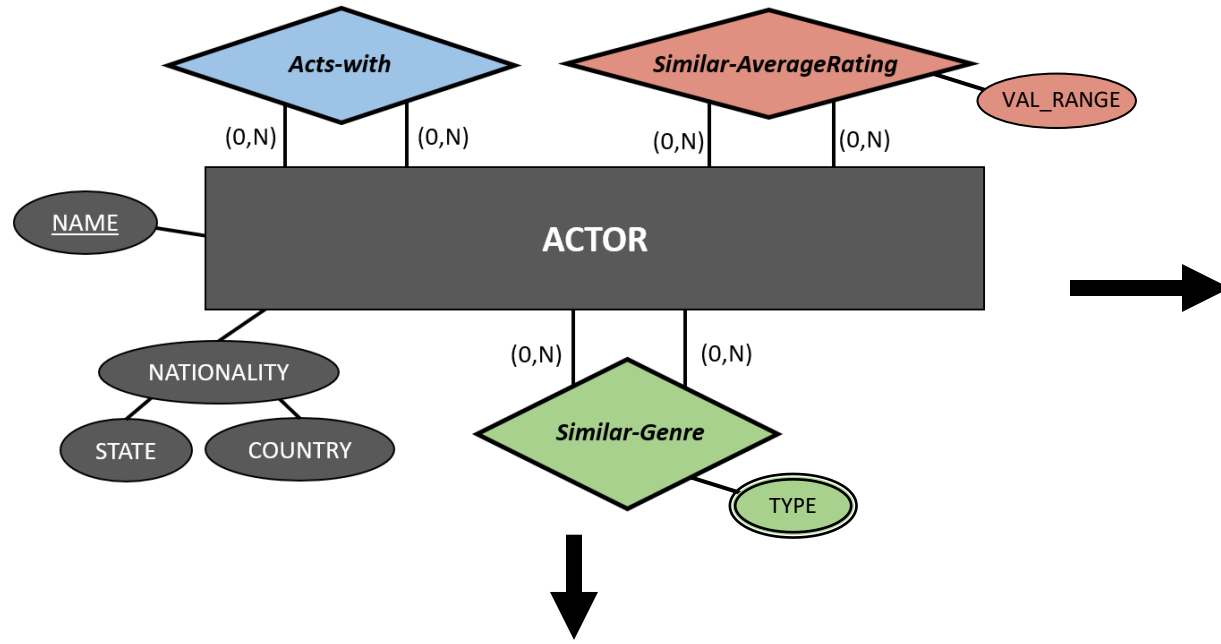
<u>AuthorID</u>	<u>YearID</u>
-----------------	---------------

### Writes

<u>AuthorID</u>	<u>PaperID</u>
-----------------	----------------

# EER Model → MLN Model: The 8 Step Algorithm

## Actor Interaction Data Set (IMDb) Modeling



Relations obtained as by product used for Drill-Down Analysis

**Actor**

<u>Name</u>	State	Country
-------------	-------	---------

**Similar-Genre\_TYPE**

<u>Actor1Name</u>	<u>Actor2Name</u>	Type
-------------------	-------------------	------

**Acts-with**

<u>Actor1Name</u>	<u>Actor2Name</u>
-------------------	-------------------

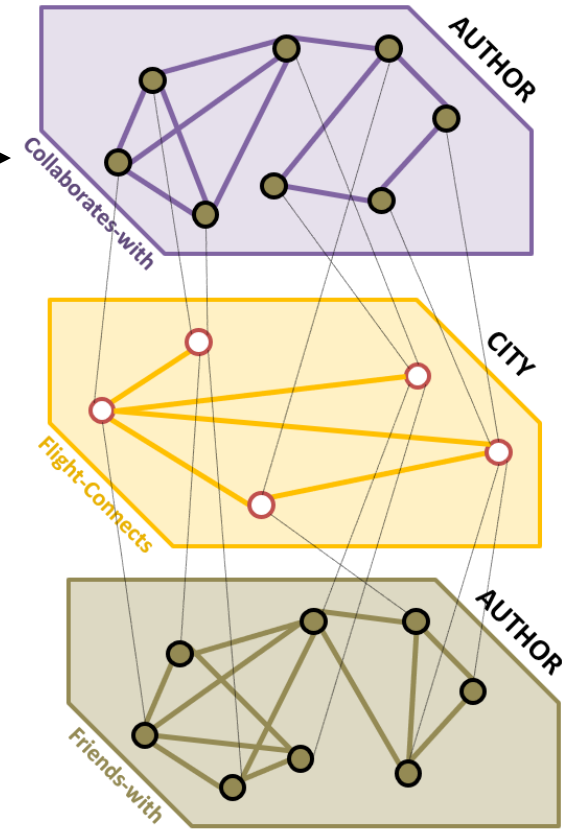
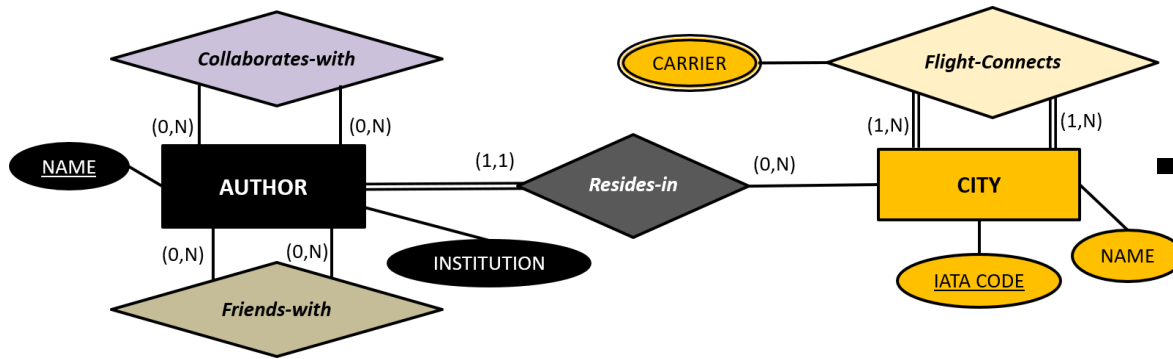
**Similar-AverageRating**

<u>Actor1Name</u>	<u>Actor2Name</u>	Val_Range
-------------------	-------------------	-----------

**Homogeneous MLN (HoMLN)**

# EER Model → MLN Model: The 8 Step Algorithm

## Author-City Interaction Data Set Modeling



**Relations obtained as by product used for Drill-Down Analysis**

### Author

<u>Name</u>	Institution	ResidenceCODE
-------------	-------------	---------------

### Friends-with

<u>Author1Name</u>	<u>Author2Name</u>
--------------------	--------------------

### Collaborates-with

<u>Author1Name</u>	<u>Author2Name</u>
--------------------	--------------------

### City

<u>IATA CODE</u>	Name
------------------	------

### Flight-Connects\_CARRIER

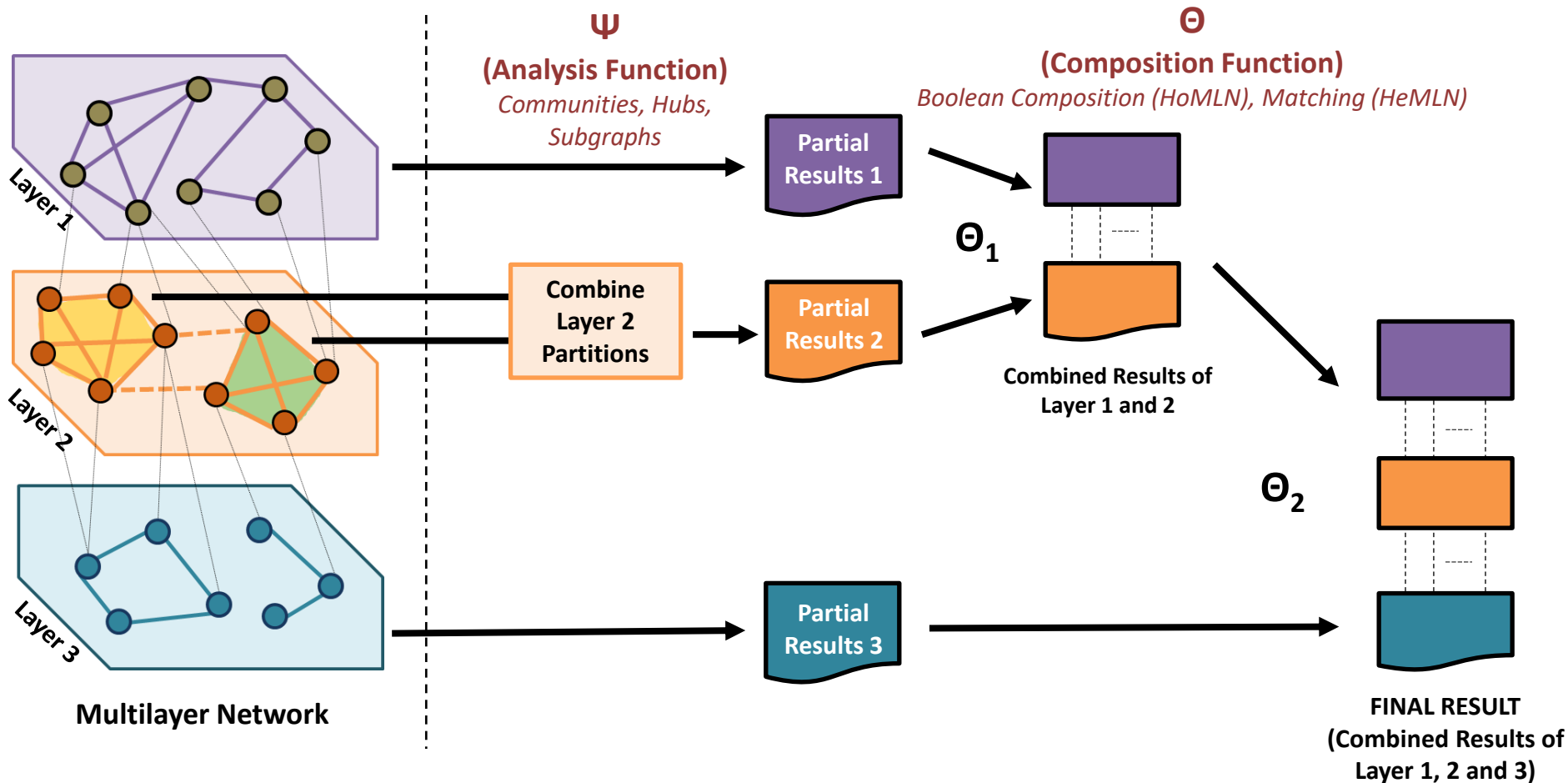
<u>City1Code</u>	<u>City2Code</u>	Carrier
------------------	------------------	---------

**Hybrid MLN (HyMLN)**



# Analysis Method: Decoupling Approach

**Divide and Conquer Approach:** Analysis function-specific partial (or intermediate) results composed systematically to fulfill objective



# Specification Mapping: Objective → MLN Expression

Objective	Mapping	
<p>Highly rated actor groups working in similar genres but have <b>not co-acted together</b> in any movie?</p>	<p>MLN Expression</p>	<p>NOT(<i>Acts-with</i>) <math>\Theta</math> <i>Similar-Genre</i> <math>\Theta</math> <i>Similar-AverageRating</i></p>
<p>HoMLN: <i>Acts-with, Similar-Genre, Similar-AverageRating</i></p>	<p><math>\Psi</math></p>	<p>Community Detection</p>
	<p><math>\Theta</math></p>	<p>Boolean AND Composition</p>
<p>For the most popular collaborators in each conference, the most active 3-year period(s)?</p>	<p>MLN Expression</p>	<p><i>Paper</i> <math>\Theta</math> <i>Author</i> <math>\Theta</math> <i>Year</i></p>
<p>HeMLN: <i>Author, Year, Paper, Review</i></p>	<p><math>\Psi</math></p>	<p>Community Detection</p>
	<p><math>\Theta</math></p>	<p>Maximal Weighted Matching</p>
<p>Best city to hold conferences of authors to maximize attendance?</p>	<p>MLN Expression</p>	<p><i>Au-Collaborates-with</i> <math>\Theta</math> <i>Au-Friends-with</i> <math>\Theta</math> <i>City</i></p>
<p>HyMLN: <i>City, Au-Collaborates-with, Au-Friends-with</i></p>	<p><math>\Psi</math></p>	<p>Community (Author), Degree Centrality (City)</p>
	<p><math>\Theta</math></p>	<p>MLN-Searching</p>

# Drill-Down Analysis: Potential Actor Collaborations

Highly rated actors working in similar genres but have *not co-acted together* in any movie

	#Vertices	#Edges in L1	#Edges in L2	#Edges in L3
<b>IMDb HoMLN</b> (For top 500 actors then repopulated with co-actors)	<b>9,485</b> (Actors)	<b>45,581</b> (Acts-with)	<b>13,945,912</b> (Similar-Genre)	<b>996,527</b> (Similar-AverageRating)

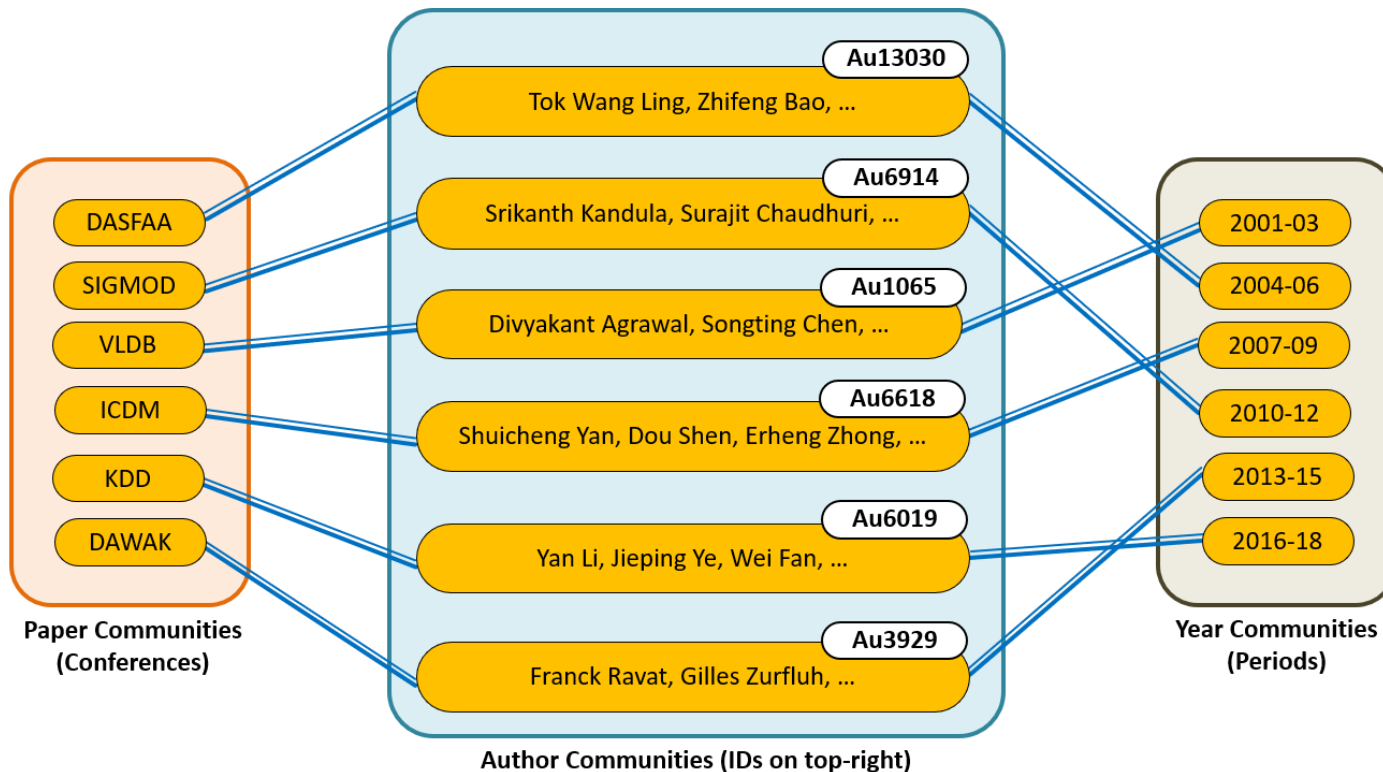
Actor/Actresses	Prominent Genres
Willem Dafoe, Russell Crowe	Action, Crime
Hilary Swank, Kate Winslet	Drama
Tom Hanks, Reese Witherspoon, Cameron Diaz	Comedy, Romance
<b>Johnny Depp, Tom Cruise</b>	<b>Adventure, Action</b>
<b>Leonardo DiCaprio, Ryan Gosling</b>	<b>Crime, Romance</b>
Nicolas Cage, Antonio Banderas	Action, Thriller
Hugh Grant, Kate Hudson, Emma Stone	Comedy, Romance

**Validating Fact:** In 2017, talks of casting **Johnny Depp** and **Tom Cruise** in pivotal roles in **Universal Studios' cinematic universe titled Dark Universe**

# Drill-Down Analysis: Research Activity Insights

For the most popular collaborators in each conference, the most active 3-year period(s)

DBLP HeMLN	Author	Paper	Year
Number of Nodes	16,918	10,326	18
Number of Edges	2,483	12,044,080	18



## Validating Facts

Most popular researchers active in different periods

**SIGMOD:** Srikanth Kandula (15188 citations)

**VLDB:** Divyakant Agrawal (23727 citations)

**ICDM:** Shuicheng Yan (52294 citations)

# Conclusions

- Proposed a **novel 8-step algorithm for MLN modeling**
  - Leveraged **EER modeling**
  - Makes the process **error-free, precise and unambiguous**
  - Aids in **drill-down analysis** of final results
- Demonstrated the **applicability on real-world applications**
- **Current Work:** Approach being used for the analysis and visualization of **spread of Covid-19 across US counties**

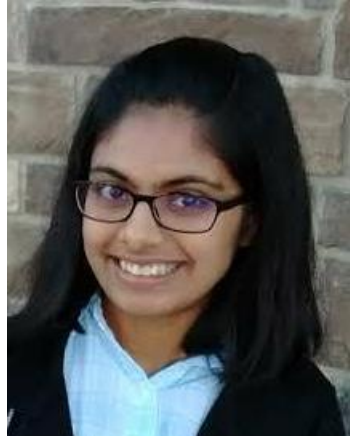
# Questions?



UNIVERSITY OF  
TEXAS  
ARLINGTON



**Sharma Chakravarthy**  
sharmac@cse.uta.edu



**Kanthi Komar**  
kanthisannappa.komar@mavs.uta.edu



**Abhishek Santra**  
abhishek.santra@mavs.uta.edu



**Sanjukta Bhowmick**  
sanjukta.bhowmick@unt.edu

Project Funded by:  National Science Foundation  
WHERE DISCOVERIES BEGIN

For more information visit:

<https://itlab.uta.edu/>



Covid-19 Analysis  
with MLNs

