Part I

Project Submission : Mock Popcorn

CSN – 351 DBMS 2017-18

Submitted By:

Ayush Chauhan 14115027

**Kunal Bansal 14116035**

Shobhit Mittal 14116061

Joydeep Das 14116032

**Shubham Agarwal 14116063**

I. Assumptions

- Video quality is considered to be a single valued attribute.
- No two movies with same title are released in the same year.
- Synopsis is unique for every movie. Movies with same title or movie sequels don't have the same synopsis.
- Each user can access his account only through one system at a time for the purpose of submitting reviews. Although, an account may be logged in through multiple systems at a time.

II. E-R Diagram
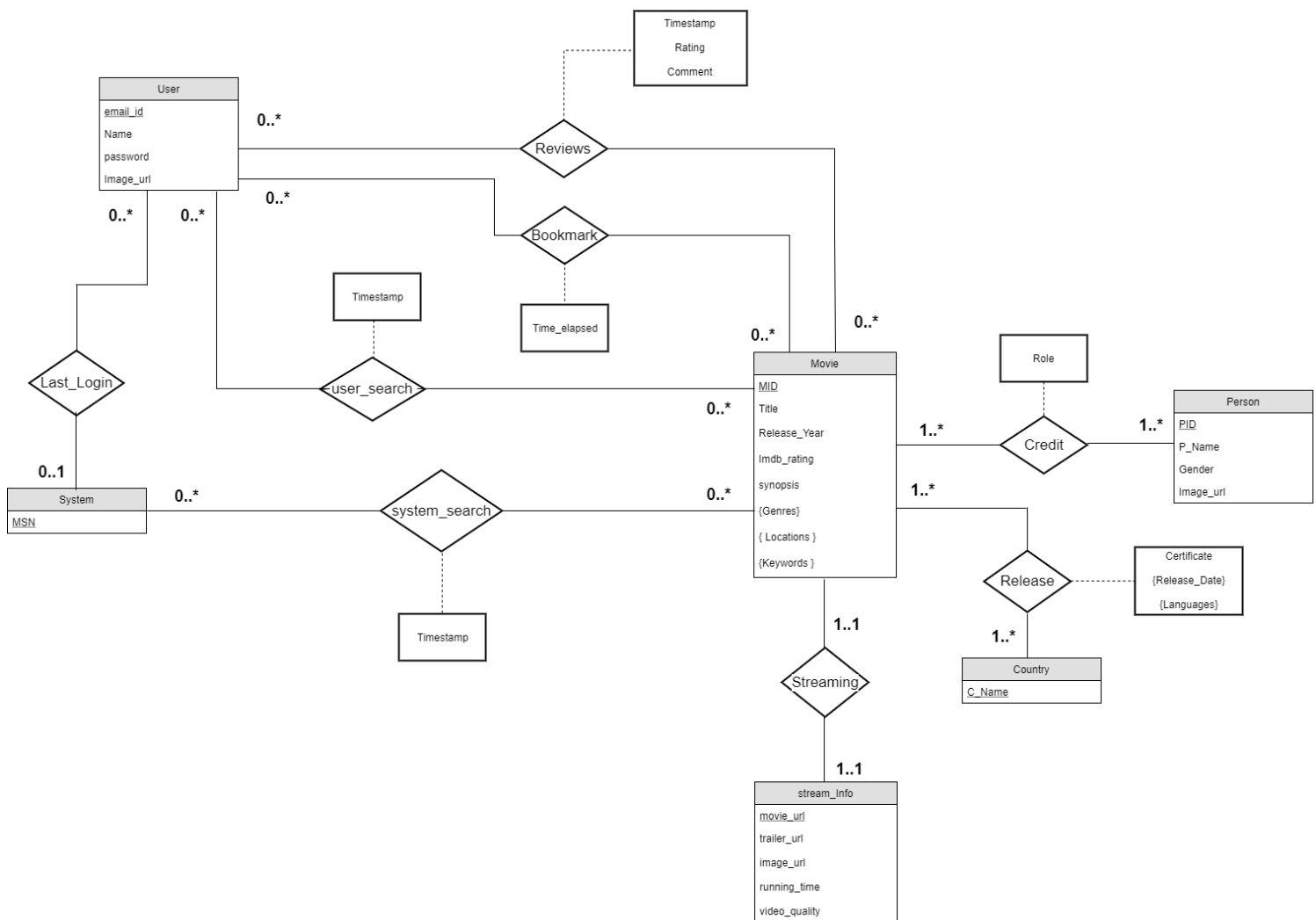
III.   Schema Diagram

Keeping in mind all the cardinality and integrity constraints, the above entities and relations are represented as tables in the following Schema diagram:
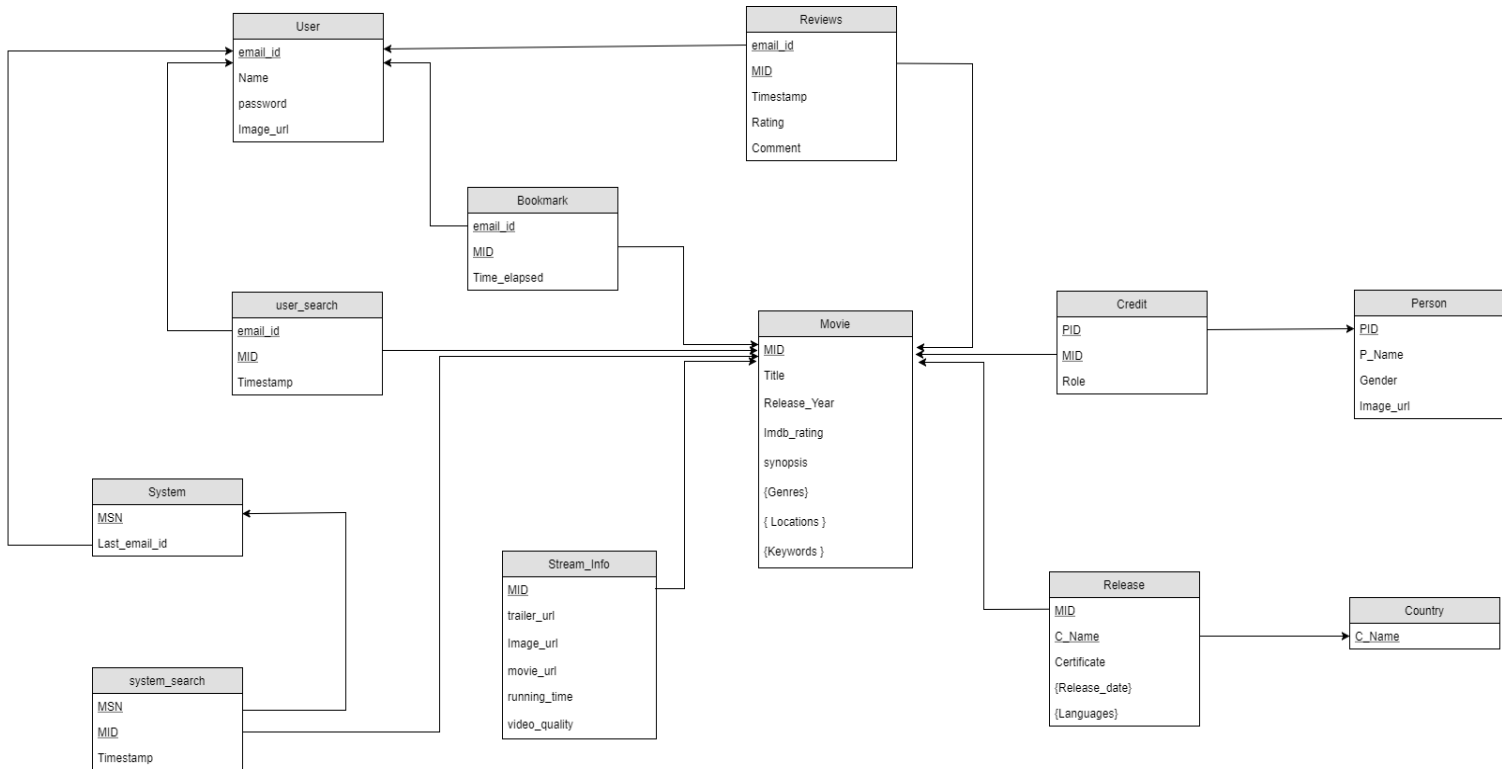
IV.   Functional Dependencies

Movie :

MID ->Title, Release_Year, Imdb_rating, Synopsis, Genres, Locations, Keywords

(Title, Release_Year) ->MID, Imdb_rating, Synopsis,Genres,Locations, Keywords

Synopsis -> MID, Title, Release_Year, Imdb_rating, Genres, Locations, Keywords

Person :

PID -> P_Name, Gender, Image_url

Image_url -> PID, P_Name, Gender

Credit :

(MID,PID) -> Role

Country :

No non-trivial FD's.

Release :

(MID, C_Name) -> Certificate, Release_Date, Languages

Stream_Info:

MID -> Movie_url, Image_url, Running_time, Video_quality

Movie_url -> MID, Trailer_url, Image_url, Running_time, Video_quality

Trailer_url -> MID, Movie_url, Image_url, Running_time, Video_url

Image_url -> MID, Movie_url, Trailer_url, Running_time, Video_quality

User:

Email_Id -> Name, Password, Image_url

Reviews:

(Email_id, MID) -> Timestamp, Rating, Comment

(Email_id, timestamp) -> MID, Rating, Comment

Bookmarks:

(Email_id, MID) -> time_elapsed

System:

MSN -> Last_email_id

User_Search:

(Email_id, MID) -> Timestamp

(Email_id, Timestamp) -> MID

System_search:

(MID, MSN) -> timestamp

(MSN, Timestamp) -> MID

V.   Minimal Cover

Movie:

<u>Step 1</u>.

MID   -> Title

MID   -> Release_Year

MID   -> Imdb_rating

MID   -> Synopsis

MID   -> Genres

MID   -> Locations

MID   -> Keywords

(Title, Release_Year)   -> MID

(Title, Release_Year)   -> Imdb_rating

(Title, Release_Year)   -> Synopsis

(Title, Release_Year)   -> Genres

(Title, Release_Year)   -> Locations

(Title, Release_Year)   -> Keywords

Synopsis   -> MID

Synopsis   ->Title

Synopsis   -> Release_Year

Synopsis   -> Imdb_rating

Synopsis   -> Genres

Synopsis   -> Locations

                    -> Keywords

Step 2:

$\{Release\_Year\}^{+} = \{ Release\_Year \}$

$\{Title\}^{+} = \{Title \}$

Hence, no reduction possible.

Step 3:

MID   -> Title

MID   -> Release_Year

MID   -> Imdb_rating

MID   -> Synopsis

MID   -> Genres

MID   -> Locations

MID   -> Keywords

(Title, Release_Year)   -> MID

Synopsis   -> MID

Person:

Step 1:

PID -> P_Name

PID -> Gender

PID -> Image_url

Image_url -> PID

Image_url -> P_Name

Image_url -> Gender

Step 2 :

No reduction possible.

Step 3:

PID -> P_Name

PID -> Gender

PID -> Image_url

Image_url -> PID

Credit:

(MID,PID) -> Role

No reductions possible.

Country:

No non-trivial FD's

Release:

(MID, C_Name)   -> Certificate

(MID, C_Name)   -> Release_Date

(MID, C_Name)   -> Languages

No reductions possible.

Stream_Info:

Step 1:

MID   -> Movie_url

MID   ->Trailer_url

MID   -> Image_url

MID   -> Running_time

MID   -> Video_quality

Movie_url   -> MID

Movie_url   -> Trailer_url

Movie_url   -> Image_url

Movie_url   -> Running_time

Movie_url   -> Video_quality

Trailer_url   -> MID

Trailer_url   -> Movie_url

Trailer_url   -> Image_url

Trailer_url   -> Running_time

Trailer_url   -> Video_quality

Image_url   -> MID

Image_url   -> Trailer_url

Image_url   -> Movie_url

Image_url   -> Running_time

Image_url   -> Video_quality

Step 2:

No reduction possible.

Step 3:

MID   -> Movie_url

MID   ->Trailer_url

MID   -> Image_url

MID   -> Running_time

MID   -> Video_quality

Movie_url   -> MID

Trailer_url   -> Movie_url

Image_url   -> Trailer_url

User:

Email_Id -> Name

Email_Id -> Password

Email_Id -> Image_url

No reductions possible.

Reviews:

<u>Step 1:</u>

(Email_id, MID)   -> Timestamp

(Email_id, MID)   -> Rating

(Email_id, MID)   -> Comment

(Email_id, timestamp) -> MID

(Email_id, timestamp) -> Rating

(Email_id, timestamp) -> Comment

<u>Step 2:</u>

No reductions possible.

<u>Step 3:</u>

(Email_id, MID) -> Timestamp

(Email_id, MID) -> Rating

(Email_id, MID) -> Comment

(Email_id, timestamp) -> MID

Bookmarks:

(Email_id, MID) -> time_elapsed

No reductions possible.

System:

MSN -> Last_email_id

No reductions possible.

User_Search:

(Email_id, MID) -> Timestamp

(Email_id, Timestamp) -> MID

No reductions possible.

System_search:

(MID, MSN) -> timestamp

(MSN, Timestamp) -> MID

No reductions possible.

VI.   Normalization

The keys, prime and nonprime attributes for each table are written before decomposition. The decomposition up to BCNF is such that all the dependencies are preserved and the decomposed tables form a lossless join.

Movie:

MID -> Title

MID -> Release_Year

MID -> Imdb_rating

MID -> Synopsis

MID -> Genres

MID -> Locations

MID -> Keywords

(Title, Release_Year) -> MID

Synopsis -> MID

Candidate keys: { (MID), (Title, Release_Year), (Synopsis) }

Prime Attributes: { MID, Title, Release_Year, Synopsis }

Non-Prime Attributes: { Imdb_rating, Genres, Locations, Keywords }

<u>1NF:</u>

Multi-valued attributes: {Genres, Locations, Keywords}

Therefore, making all attributes atomic for 1NF, the decomposed tables are:

Movie: {MID, Title, Release_Year, Imdb_rating, Synopsis}

Keys: { (MID), (Title, Release_Year), (Synopsis) }

Movie_Genres: {MID, Genres}

Keys: { (MID, Genres) }

Movie_Locations: {MID, Locations}

Keys: { (MID, Locations) }

Movie_keywords: {MID, Keywords}

Keys: { (MID, Keywords) }

2NF:

Movie:

The only Non-prime attribute (*Imdb_rating*) is fully functionally dependent on all the candidate keys of Movie. Therefore, the table is already in 2NF.

Movie_Genres:

No non-prime attribute.

Movie_Locations:

No non-prime attribute.

Movie_keywords:

No non-prime attribute.

3NF:

Movie:

No transitive dependency.

Movie_Genres:

No non-prime attribute.

Movie_Locations:

No non-prime attribute.

Movie_keywords:

No non-prime attribute.

BCNF:

Movie:

For every FD: X->Y, X is a superkey, hence already in BCNF.

Movie_Genres:

Already in BCNF.

Movie_Locations:

Already in BCNF.

Movie_keywords:

Already in BCNF.

4NF:

Movie:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

Movie_Genres, Movie_Locations, Movie_keywords:

No non-trivial MVD, hence already in 4NF.

5NF:

Movie:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

Movie_Genres, Movie_Locations, Movie_keywords:

No non-trivial JD, hence already in 5NF.

Person:

PID -> P_Name

PID -> Gender

PID -> Image_url

Image_url -> PID

Candidate keys: { (PID), (Image_url) }

Prime Attributes: { PID, Image_url }

Non-Prime Attributes: { P_Name, Gender }

1NF:

No multivalued attribute, already in 1NF.

2NF:

All the candidate keys are single attribute keys; therefore the table is in 2NF.

3NF:

None of the non-prime attributes have a transitive dependency on any of the keys.

BCNF:

For every FD: X->Y, X is a superkey, hence already in BCNF.

4NF:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

5NF:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

Credit :

(MID,PID) -> Role

Candidate keys: { (MID,PID) }

Prime Attributes: { MID, PID}

Non-Prime Attributes: { Role }

1NF:

No multivalued attribute, already in 1NF.

2NF:

The non-prime attribute Role is fully functionally dependent on the key (*MID, PID*).

3NF:

No transitive dependency from the key to any non-prime attribute.

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

No MVD from a non-superkey tuple; therefore, already in 4NF.

<u>5NF:</u>

No non-trivial JD exists in this table; hence, already in 5NF.

Country:

This table contains only one attribute, *C_Name*. Thus, this table satisfies conditions for all the normalised forms, i.e. <u>1NF</u>, <u>2NF</u>, <u>3NF</u>, <u>BCNF</u>, <u>4NF</u>, <u>5NF</u>.

Release:

(MID, C_Name)   -> Certificate

(MID, C_Name)   -> Release_Date

(MID, C_Name)   -> Languages

Candidate keys: { (MID, C_Name) }

Prime Attributes: { MID, C_Name}

Non-Prime Attributes: { Certificate, Release_Date, Languages}

<u>1NF:</u>

Multi-valued attributes: { Release_Date, Languages}

Therefore, making all attributes atomic for 1NF, the decomposed tables are:

Movie_Certificates: { MID, C_Name, Certificate}

Keys : { (MID, C_Name) }

Movie_dates: { MID, C_Name, Release_Date}

Keys: { (MID, C_Name, Release_Date) }

Movie_Languages: { MID, C_Name, Languages}

Keys: { (MID, C_Name, Languages) }

<u>2NF:</u>

Movie_Certificates:

Non-prime attribute (*Certificate*) is fully functionally dependent on the key (*MID, C_Name*).

Movie_dates:

No non-prime attribute, already in 2NF.

Movie_Languages:

No non-prime attribute, already in 2NF.

3NF:

Movie_Certificates:

No transitive dependency from the key to non-prime attribute, hence already in 3NF.

Movie_dates:

No non-prime attribute, already in 3NF.

Movie_Languages:

No non-prime attribute, already in 3NF.

BCNF:

Movie_Certificates:

The LHS of the only FD is a superkey, hence already in BCNF.

Movie_dates:

No non-trivial FD, hence already in BCNF.

Movie_Languages:

No non-trivial FD, hence already in BCNF.

4NF:

No non-trivial MVD in any of the decomposed tables, hence all of them are in 4NF.

5NF:

No non-trivial JD in any of the decomposed tables, hence all of them are in 5NF.

Stream_Info:

MID   -> Movie_url

MID   ->Trailer_url

MID   -> Image_url

MID   -> Running_time

MID   -> Video_quality

Movie_url -> MID

Trailer_url -> Movie_url

Image_url -> Trailer_url

Candidate keys: { (MID), (Movie_url), (Trailer_url), (Image_url) }

Prime Attributes: { MID, Movie_url, Trailer_url, Image_url }

Non-Prime Attributes: { Running_time, Video_quality}

<u>1NF:</u>

No multivalued attribute, already in 1NF.

<u>2NF:</u>

All the candidate keys are single attribute keys, therefore the table is in 2NF.

<u>3NF:</u>

There is no transitive dependency from the key to any non-prime attribute.

<u>BCNF:</u>

For every FD: X->Y, X is a superkey, already in BCNF.

<u>4NF:</u>

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

<u>5NF:</u>

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

User:

Email_Id -> Name

Email_Id -> Password

Email_Id -> Image_url

Candidate keys: { Email_id }

Prime Attributes: { Email_id }

Non_prime Attributes: { Name, Password, Image_url }

1NF:

No multivalued attribute, already in 1NF.

2NF:

The only candidate key (*Email_id*) is a single attribute key; therefore the table is in 2NF.

3NF:

No transitive dependency exists in the table.

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

No non-trivial MVD, hence already in 4NF.

5NF:

No non-trivial JD, hence already in 5NF.

Reviews:

(Email_id, MID) -> Timestamp

(Email_id, MID) -> Rating

(Email_id, MID)   -> Comment

(Email_id, timestamp) -> MID

Candidate keys: { (MID, Email_id), (Email_id, Timestamp) }

Prime Attributes: { MID, Email_id, Timestamp }

Non_prime Attributes: { Rating, Comment }

1NF:

No multivalued attribute, already in 1NF.

2NF:

All non-prime attributes are fully functionally dependent on both the keys of the table.

3NF:

No transitive dependency.

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

5NF:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

Bookmarks:

(Email_id, MID) -> time_elapsed

Candidate keys: { (MID, Email_id) }

Prime Attributes: { MID, Email_id }

Non_prime Attributes: { time_elapsed }

1NF:

No multivalued attribute, already in 1NF.

2NF:

The only non-prime attribute (*time_elapsed*) are fully functionally dependent on the key.

3NF:

No transitive dependency. .

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

5NF:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

System:

MSN -> Last_email_id

Candidate keys: { MSN }

Prime Attributes: { MSN }

Non_prime Attributes: { Last_email_id }

1NF:

No multivalued attribute, already in 1NF.

2NF:

The only key of the table, *MSN* is a single attribute key. Thus, the table is in 2NF.

3NF:

The table has only 1 FD, therefore it is in 3NF.

BCNF:

For the only FD, *MSN -> Last_email_id*, *MSN* is a superkey; hence, already in BCNF.

4NF:

No non-trivial MVD exists in the table; hence, already in 4NF.

5NF:

No non-trivial JD exists in the table; hence, already in 5NF.

User_Search:

(Email_id, MID) -> Timestamp

(Email_id, Timestamp) -> MID

Candidate keys: { (MID, Email_id), (Email_id, Timestamp) }

Prime Attributes: { MID, Email_id, Timestamp }

Non_prime Attributes: None

1NF:

No multivalued attribute, already in 1NF.

2NF:

No non-prime attribute, already in 2NF.

3NF:

No non-prime attribute, already in 3NF.

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

5NF:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

System_search:

(MID, MSN) -> timestamp

(MSN, Timestamp) -> MID.

Candidate keys: { (MID, MSN), (MSN, Timestamp) }

Prime Attributes: { MID, MSN, Timestamp }

Non_prime Attributes: None

1NF:

No multivalued attribute, already in 1NF.

2NF:

No non-prime attribute, already in 2NF.

3NF:

No non-prime attribute, already in 3NF.

BCNF:

For every FD: X->Y, X is a superkey, already in BCNF.

4NF:

There is no non-trivial MVD such that the LHS is not a superkey. Hence, the table is already in 4NF.

5NF:

Amongst all the non-trivial JDs that can be formed, every decomposed table turns out to be a superkey of the original table. Hence, the table is already in 5NF.

VII.   Final Normalized Tables

Movie: {MID, Title, Release_Year, Imdb_rating, Synopsis}

Movie_Genres: {MID, Genres}

Movie_Locations: {MID, Locations}

Movie_keywords: {MID, Keywords}

Person: {PID, P_Name, Gender, Image_url}

Credit: {MID, PID, Role}

Country: {C_Name}

Movie_Certificates: { MID, C_Name, Certificate}

Movie_dates: { MID, C_Name, Release_Date}

Movie_Languages: { MID, C_Name, Languages}

Stream_Info: {MID, Movie_url, Trailer_url, Image_url, Running_time, Video_quality}

User: {Email_Id, Name, Password, Image_url}

Reviews: {Email_id, MID, Timestamp, Rating, Comment}

Bookmarks: {Email_id, MID, time_elapsed}

System: {MSN, Last_email_id}

User_search: {Email_id, MID, Timestamp}

System_search: {MSN, MID, Timestamp}