



UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CdL Ingegneria Informatica

Documentazione Progetto Basi di Dati

Realizzato da

Marranini Lorenzo Moni Lorenzo

ANNO ACCADEMICO 2022/2023

Abstract

Il seguente documento tratta la documentazione relativa alla progettazione del sistema informativo automatizzato della piattaforma FilmSphere, un portale dedicato alla memorizzazione e streaming di contenuti multimediali.

Si vogliono memorizzare i dati relativi ai film presenti sulla piattaforma, ognuno dei quali sarà associato a molteplici contenuti, disponibili in formati e codifiche differenti. Si vogliono memorizzare i dati relativi agli utenti registrati, agli abbonamenti sottoscritti e alle connessioni effettuate dagli utenti attraverso i propri dispositivi. Si vogliono inoltre memorizzare i dati relativi alla Content Delivery Network come il carico attuale dei server, le erogazioni dei contenuti, e lo stato di archiviazione dei server. Il database deve inoltre mettere a disposizione delle funzionalità analytics in grado di fornire delle informazioni utili al miglioramento dell'esperienza utente e dell'efficienza della CDN.

Per garantire una maggiore chiarezza del glossario, della progettazione concettuale e del modello relazionale è stato scelto di suddividere il progetto in 5 aree tematiche, ognuna contraddistinta da un diverso colore, seguendo la divisione presentata nella documentazione originale:

- Area Contenuti
- Area Formati
- Area Clienti
- Area Streaming
- Area Analytics

La seguente documentazione presenta tutte le fasi relative alla progettazione e alla realizzazione del sistema informativo automatizzato, ponendo particolare attenzione sulle motivazioni relative alle scelte progettuali effettuate durante la realizzazione del servizio.

Indice

1 Glossario dei Termini	9
I Progettazione Concettuale	13
2 Area Contenuti	14
2.1 Sezione Diagramma Interessato	14
2.2 Dizionario Dei Dati	15
2.3 Descrizione Area	17
2.3.1 Film	17
2.3.2 Cast e Premiazioni	17
2.3.3 Rating e Recensioni	18
3 Area Formati	19
3.1 Sezione Diagramma Interessato	19
3.2 Dizionario Dei Dati	20
3.3 Descrizione Area	20
3.3.1 Codec e Formati	20
3.3.2 Restrizioni Formati	21
4 Area Clienti	22
4.1 Sezione Diagramma Interessato	22
4.2 Dizionario Dei Dati	23
4.3 Descrizione Area	24
4.3.1 Clienti e Abbonamenti	24
4.3.2 Connessioni e Dispositivi	25
4.3.3 Fatturazioni e Pagamenti	25
4.3.4 Sistema di Raccomandazione Contenuti	26
5 Area Streaming	29
5.1 Sezione Diagramma Interessato	29
5.2 Dizionario Dei Dati	30
5.3 Descrizione Area	30
5.3.1 Server e Memorizzazione	31
5.3.2 Fruizione ed Erogazione	31
5.3.3 Caching	32
5.3.4 Richiesta Contenuti e Scelta del Server	32

6 Area Analytics	34
6.1 Sezione Diagramma Interessato	34
6.2 Dizionario Dei Dati	35
6.3 Descrizione Area	35
6.3.1 Classifiche	36
6.3.2 Bilanciamento del carico	36
II Ristrutturazione	37
7 Eliminazione Generalizzazioni	38
7.1 Generalizzazione di Artista	38
7.2 Generalizzazione di Premio	39
7.3 Generalizzazione di Codec	40
8 Eliminazione Attributi Composti	42
8.1 Attributo composto Periodo	42
8.2 Attributo composto Posizione	42
III Analisi Prestazionale	43
9 Tavola dei Volumi	44
10 Operazioni sui Dati	49
10.1 Registrazione Utente e Creazione Abbonamento	49
10.1.1 Sezione Diagramma Interessato	49
10.1.2 Tavola dei Volumi Interessati	50
10.1.3 Tavola degli Accessi	50
10.2 Login alla Piattaforma	52
10.2.1 Sezione Diagramma Interessato	52
10.2.2 Tavola dei Volumi Interessati	53
10.2.3 Tavola degli Accessi	53
10.3 Emissione Fatture	54
10.3.1 Sezione Diagramma Interessato	54
10.3.2 Tavola dei Volumi Interessati	55
10.3.3 Tavola degli Accessi	55
10.3.4 Valutazione della Ridondanza Costo Mensile in Abbonamento	56
10.4 Richiesta Visualizzazione Contenuto Video	59
10.4.1 Sezione Diagramma Interessato	59
10.4.2 Tavola dei Volumi Interessati	60
10.4.3 Tavola degli Accessi	61
10.4.4 Valutazione Ridondanze Memoria Usata e Banda Usata in Server	63
10.5 Aggiornamento Rating e Classifica per Genere	68
10.5.1 Sezione Diagramma Interessato	68

10.5.2 Tavola dei Volumi Interessati	69
10.5.3 Tavola degli Accessi	70
10.5.4 Valutazione delle Ridondanze per Aggiornamento Rating	71
10.6 Server più utilizzato per l'Erogazione di un Film	75
10.6.1 Sezione Diagramma Interessato	75
10.6.2 Tavola dei Volumi Interessati	76
10.6.3 Tavola degli Accessi	76
10.7 Stati con maggiori Film premiati	77
10.7.1 Sezione Diagramma Interessato	77
10.7.2 Tavola dei Volumi Interessati	78
10.7.3 Tavola degli Accessi	78
10.8 Distribuzione delle visualizzazioni di un Film	79
10.8.1 Sezione Diagramma Interessato	79
10.8.2 Tavola dei Volumi Interessati	80
10.8.3 Tavola degli Accessi	80
IV Progettazione Logica	82
11 Descrizione Schema Logico	83
11.1 Schema Logico	83
11.2 Vincoli di Integrità	86
11.2.1 Vincoli di Integrità Referenziale	86
11.2.2 Vincoli di Integrità di Tupla	89
11.2.3 Vincoli di Integrità Generici	91
11.3 Analisi Dipendenze Funzionali e Normalizzazione	93
12 Funzioni Analytics	97
12.1 Classifiche	97
12.1.1 Classifica per Stato	97
12.1.2 Classifica per Pacchetto	98
12.1.3 Classifica per Genere	99
12.2 Bilanciamento del Carico	100
12.2.1 Gestione Excessive Load	100
12.2.2 Gestione Excessive Memory	101
12.2.3 Gestione Bilanciamento Carico	102
12.3 Riprendi Visualizzazione	103
12.4 Statistiche Lingue	104

Glossario dei Termini

Affinché l'utente possa usufruire del database al pieno delle sue potenzialità e senza frain-tenderne i contenuti, di seguito è presente un glossario dei termini ricorrenti e rilevanti, forniti ognuno di un'accurata descrizione, che ne esplica il significato ai fini del database.

Termine	Descrizione	Sinonimi	Collegamenti
Area Generale			
Stato	Organizzazione politica e giuridica di una comunità stabilmente stanziata su un territorio.	Nazione, Zona Geografica	Continente
Area Contenuti			
Film	Opera cinematografica fruibile dagli utenti in diverse lingue e formati audio/video.	Opera Cinematografica	Rating, Artista, Premio
Rating	Valutazione di un particolare film determinato da un algoritmo che considera diversi fattori.	Valutazione	Film, Utente, Critico, Premio
Artista	Persona che ha preso parte attiva alla realizzazione di un film.		Film, Premio
Attore	Artista che fa parte del cast di uno o più film.		Film, Artista
Regista	Artista che svolge il ruolo di regista in uno o più film.		Film, Artista
Critico	Persona che si occupa per mestiere di recensire contenuti video.	Recensore	Film, Rating
Premio	Riconoscimento che viene assegnato ad un contenuto video o alla performance di un artista.	Riconoscimento	Film, Artista
Area Clienti			
Utente	Persona fisica registrata alla piattaforma.	Cliente	Abbonamento

Connessione	Collegamento tra un singolo utente e la piattaforma attraverso uno specifico dispositivo e area geografica.	Accesso, Login	Utente, Erogazione
Abbonamento	Sottoscrizione alla piattaforma, associata ad un utente, che garantisce la fruizione dei contenuti fornendo sottoinsiemi di funzionalità.	Sottoscrizione	Utente, Fatturazione
Pacchetto	Piano di sottoscrizione alla piattaforma che mette a disposizione determinati contenuti e funzionalità distinte.	Piano	Utente, Fatturazione
Fatturazione	Informazione relativa al pagamento del piano di abbonamento sottoscritto da un utente.		Utente, Abbonamento
Carta di Pagamento	Carta di credito o debito che permette ad un determinato utente di saldare le fatture relative al servizio.	Carta di Credito, Carta di Debito	Utente, Abbonamento
Raccomandazione	Funzionalità del sistema in grado di generare suggerimenti in base alle caratteristiche dell'utente e dei contenuti visualizzati.		Utente, Film
Dispositivo	Device fisico utilizzato per stabilire una connessione con il servizio.	Device	Connessione
Compatibilità	Valore che definisce il grado di affinità tra un'opera cinematografica e un utente.	Affinità	Rating, Utente, Film
Area Formati			
Contenuto	File fisico relativo ad un film che risiede su uno o più server della CDN relativo ad un film.	File	Film, Archiviazione
Formato	Tipologia di formato di archiviazione di un determinato file. Un formato può essere disponibile in più codec.		Film, Codec

Codec	Algoritmo di codifica e decodifica dei dati di un formato audio/video. Un codec è sempre associato ad un formato.	Codifica	Formato
Area Streaming			
Streaming	Flusso di dati video trasmessi da una sorgente a una o più destinazioni tramite una rete telematica.	Trasmissione	Server, CDN, Contenuto, Erogazione
Content Delivery Network	Rete di distribuzione di contenuti formata da più server dislocati in diverse località geografiche.	CDN, Rete distribuita	Server
Server	Macchina fisica connessa in rete in grado di fornire contenuti attraverso flussi streaming.		CDN
Caching	Metodo che permette di ridurre il buffering e la latenza attraverso la memorizzazione di contenuti su server intermediari.		CDN, Server, Film
Erogazione	Flusso di streaming fornito da un server che permette la visualizzazione di un contenuto da parte di un utente.	Visualizzazione	Server, CDN
Collegamento	Connessione tra un server e un'erogazione per lo streaming di un contenuto multimediale.		Server, CDN, Contenuto
Bilanciamento del Carico	Metodo che permette di distribuire le erogazioni degli utenti su più server.	Load Balancing	Server, CDN
Sovraccarico	Stato in cui un server non è più in grado di erogare ulteriori contenuti a causa dell'assenza di banda disponibile.		Server, CDN
Restrizione	Limitazione imposta allo streaming di un determinato contenuto.	Limitazione	Contenuto, Stato, Pacchetto
Area Analytics			
Server Log	Informazione relativa ad un server della CDN.		Server, CDN

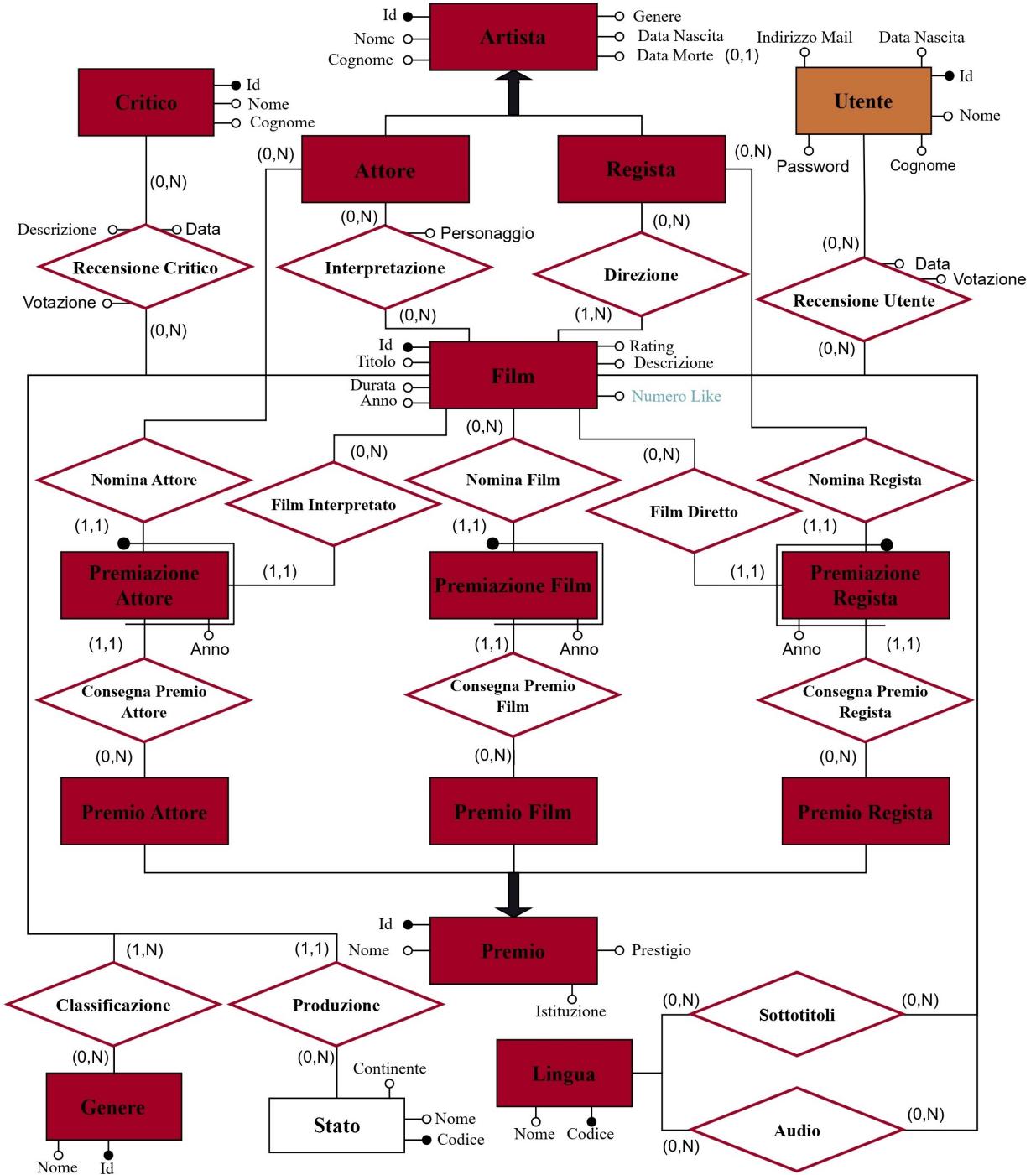
Classifica	Graduatoria stilata in base a specificati parametri.	Graduatoria	Film, Stato, Generi, Pacchetto
------------	--	-------------	--------------------------------

Parte I

Progettazione Concettuale

Area Contenuti

2.1 Sezione Diagramma Interessato



2.2 Dizionario Dei Dati

Entità	Descrizione	Attributi	Identifieri
Artista	Persona che lavora nell'ambito cinematografico.	Id, Nome, Cognome, Data di Nascita, Data di Morte	Id
Regista	Artista che svolge il ruolo di regista.		Id(Artista)
Attore	Artista che interpreta un ruolo in un film.		Id(Artista)
Film	Opera cinematografica.	Id, Titolo, Descrizione, Durata, Anno di produzione, Rating, Numero Like	Id
Genere	Genere relativo ad un'opera cinematografica.	Id, Nome	Id
Lingua	Lingua in cui è disponibile un determinato contenuto video.	Codice, Nome	Codice
Critico	Persona che si occupa di esaminare e giudicare un contenuto video.	Id, Nome, Cognome	Id
Premio	Premio relativo ad un'opera cinematografica o ad un artista.	Id, Nome, Anno, Istituzione, Prestigio	Id
Premio Film	Premio associato ad un Film.		Id(Premio)
Premio Attore	Premio associato ad un Attore.		Id(Premio)
Premio Regista	Premio associato ad un Regista.		Id(Premio)
Premiazione Attore	Premiazione di un attore per il ruolo interpretato in un'opera cinematografica.	Anno	Id(Artista), Id(Premio Attore), Id(Film), Anno
Premiazione Regista	Premiazione di un regista per la direzione di un'opera cinematografica.	Anno	Id(Artista), Id(Premio Regista), Id(Film), Anno
Premiazione Film	Premiazione di un'opera cinematografica.	Anno	Id(Premio Film), Id(Film), Anno

Relationship	Descrizione	Attributi	Entità coinvolte
Interpretazione	Lavoro svolto da un attore per la realizzazione di un film.	Personaggio	Attore(0,N) Film(0,N)
Direzione	Lavoro svolto da un regista per la realizzazione di un film.		Regista(0,N) Film(1,N)
Classificazione	Appartenenza di un film ad un determinato genere.		Genere(0,N) Film(1,N)
Recensione Critico	Recensione pubblicata da un critico cinematografico.	Descrizione, Data, Votazione	Critico(0,N) Film(0,N)
Recensione Utente	Recensione pubblicata da un utente della piattaforma.	Data, Votazione	Utente(0,N) Film(0,N)
Produzione	Stato di produzione di un film.		Stato(0,N) Film(1,1)
Audio	Disponibilità di una traccia audio per un determinato film in una lingua.		Lingua(0,N) Film(0,N)
Sottotitoli	Disponibilità di sottotitoli per un determinato film in una lingua.		Lingua(0,N) Film(0,N)
Nomina Attore	Nomina di un attore per un premio.		Attore(0,N) Premiazione Attore(1,1)
Nomina Regista	Nomina di un regista per un premio.		Regista(0,N) Premiazione Regista(1,1)
Nomina Film	Nomina di un film per un premio.		Film(0,N) Premiazione Film(1,1)
Film Interpretato	Film relativo alla premiazione di un attore.		Film(0,N) Premiazione Attore(1,1)
Film Diretto	Film relativo alla premiazione di un regista.		Film(0,N) Premiazione Regista(1,1)

Consegna Premio Attore	Premio relativo alla premiazione di un attore.		Premio Attore(0,N) Premiazione Attore(1,1)
Consegna Premio Regista	Premio relativo alla premiazione di un regista.		Premio Regista(0,N) Premiazione Regista(1,1)
Consegna Premio Film	Premio relativo alla premiazione di un film.		Premio Film(0,N) Premiazione Film(1,1)

2.3 Descrizione Area

L'area contenuti è la zona del database dedicata alla memorizzazione delle informazioni relative ai contenuti video presenti sulla piattaforma.

2.3.1 Film

Un film è contraddistinto da un identificatore univoco e presenta informazioni quali il titolo, una breve descrizione, l'anno di rilascio, la durata in minuti e i generi a cui tale film appartiene. Per ogni film si hanno inoltre molteplici lingue disponibili, sia per quanto riguarda le tracce audio, che i sottotitoli. Tale funzionalità è stata implementata attraverso la creazione di due relationship molti a molti tra l'entità *Film* e l'entità *Lingua*. Per quanto riguarda la relationship *Audio* è stata scelta la cardinalità opzionale per l'entità *Film* a causa del fatto che un film potrebbe non avere tracce audio, e dunque essere muto.

2.3.2 Cast e Premiazioni

Per ogni film è disponibile il cast che ha lavorato alla produzione e i relativi premi che sono stati vinti in seguito alla realizzazione di tale film.

Ogni persona che lavora alla produzione di un film viene descritta come istanza dell'entità *Artista* che possiede due specializzazioni *Attore* e *Regista*.

Il sistema di premiazioni è implementato attraverso tre entità: *Premiazione Attore*, *Premiazione Regista*, che rappresentano rispettivamente l'attribuzione di un determinato premio ad un attore o regista per la realizzazione di un film in un anno specifico, e *Premiazione Film* che rappresenta la premiazione di un film in quanto tale.

Deve essere possibile creare un'istanza di *Premiazione Attore* e *Premiazione Regista* rispettivamente se tale attore o regista hanno partecipato alla realizzazione del film.

Ogni *Premio* ha un *nome*, un *istituzione* e un *prestigio* che viene rappresentato da un valore massimo di 5 punti.

2.3.3 Rating e Recensioni

Il sistema presenta un sistema di recensioni che permette al pubblico della piattaforma di esprimere un giudizio sui contenuti visionati. Sono state implementate due tipologie di recensioni: *Recensione Critico* e *Recensione Utente*. Il *Critico* è una figura che svolge la professione di recensire contenuti video e di conseguenza la scelta è ricaduta sull'implementazione di una recensione che permette di inserire sia una votazione in decimi, sia un messaggio aggiuntivo per esprimere un giudizio articolato sul contenuto.

Per quanto riguarda le recensioni degli utenti *FilmSphere* ha optato per un sistema di valutazione binario, tipico delle piattaforme di streaming, che permette di ridurre al minimo la frizione correlata alla formulazione di una recensione più complessa, e di conseguenza consente la raccolta di un quantitativo di recensioni maggiore, aspetto cruciale per quanto riguarda il sistema di raccomandazioni della piattaforma. L'utente per poter recensire un film deve aver visionato per almeno il 25% della sua durata uno dei contenuti associati a quel film.

Ogni film inoltre possiede un rating, espresso in centesimi, che viene utilizzato dal sistema di raccomandazioni per generare classifiche basate sul genere. Il rating di un determinato film dipende da vari fattori quali: recensioni attribuite al film dagli utenti, recensioni attribuite al film dai critici, premi ottenuti dal film, premi ottenuti dal cast per la realizzazione del film e il tempo di visualizzazione medio del film.

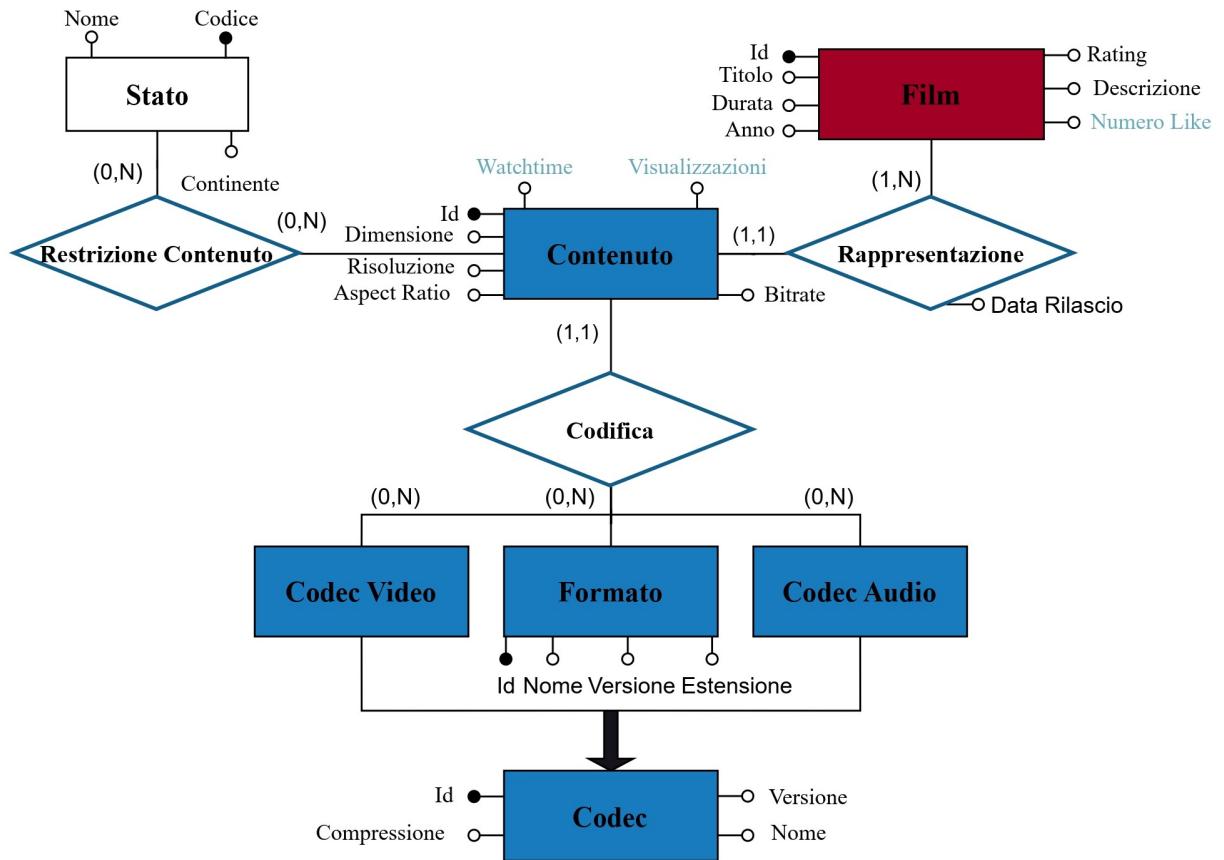
Il calcolo del rating avverrà attraverso una media pesata dei valori sopra descritti, dove ognuno dei quali avrà un peso dato dalla rilevanza dell'aspetto considerato:

- Percentuale recensioni positive utenti: Numero Like / Totale Visualizzazioni. **Peso:** 20%
- Percentuale recensioni critici: Somma valutazioni critici / (Totale recensioni critici * 10). **Peso:** 25%
- Premi ottenuti dal Film: Somma prestigio premi ricevuti dal film / (Totale premi ricevuti dal film * 5). **Peso:** 15%
- Premi ottenuti dal Cast: Somma prestigio premi ricevuti dal cast / (Totale premi ricevuti dal cast * 5). **Peso:** 10%
- Il tempo di visualizzazione medio del film: Somma durate visualizzazione del film / (Totale visualizzazioni * Durata film). **Peso:** 30%

$$\text{Rating} = \sum \text{Fattore} \cdot \text{Peso}$$

Area Formati

3.1 Sezione Diagramma Interessato



3.2 Dizionario Dei Dati

Entità	Descrizione	Attributi	Identifieri
Formato	Tipo di formato di un file.	Id, Nome, Versione, Estensione	Id
Contenuto	Copia fisica di un film che viene immagazzinato su uno o più server e ha un relativo formato e codifica.	Id, Dimensione, Risoluzione, Aspect Ratio, Bitrate, Watchtime, Visualizzazioni	Id
Codec	Tipo di codifica di un file.	Id, Nome, Compressione, Versione	Id
Codec Audio	Tipo di codifica audio.		Id(Codec)
Codec Video	Tipo di codifica video.		Id(Codec)

Relationship	Descrizione	Attributi	Entità coinvolte
Rappresentazione	Contenuto video relativo ad un determinato film.	Data Rilascio	Contenuto(1,1) Film(1,N)
Codifica	Codifica relativa a un determinato contenuto.		Contenuto(1,1) Formato(0,N) CodecAudio(0,N) CodecVideo(0,N)
Restrizione Contenuto	Possibile restrizione geografica di un contenuto.		Contenuto(0,N) Stato(0,N)

3.3 Descrizione Area

L'area formati è la zona del database dedicata alla memorizzazione delle informazioni relative ai formati disponibili per ogni film della piattaforma.

3.3.1 Codec e Format

Ogni film può essere disponibile in più formati sulla piattaforma. Ogni istanza dell'entità *Contenuto* rappresenta una copia fisica di un determinato film che risiede su uno o più server della CDN di FilmSphere. Ogni *Contenuto* è contraddistinto da una dimensione fisica espressa in GB, un bitrate espresso in Mbit/s, una risoluzione, un aspect ratio, un formato, un codec audio e un codec video.

Inoltre ogni contenuto relativo ad un determinato film possiede tutte le tracce audio e i sottotitoli disponibili per tale film.

Ogni *Formato* è rappresentato da un *nome*, una *versione* e un'*estensione*, mentre un *Codec* è contraddistinto da un *nome*, da un tipo di *compressione* e presenta due specializzazioni *Codec Audio* e *Codec Video*.

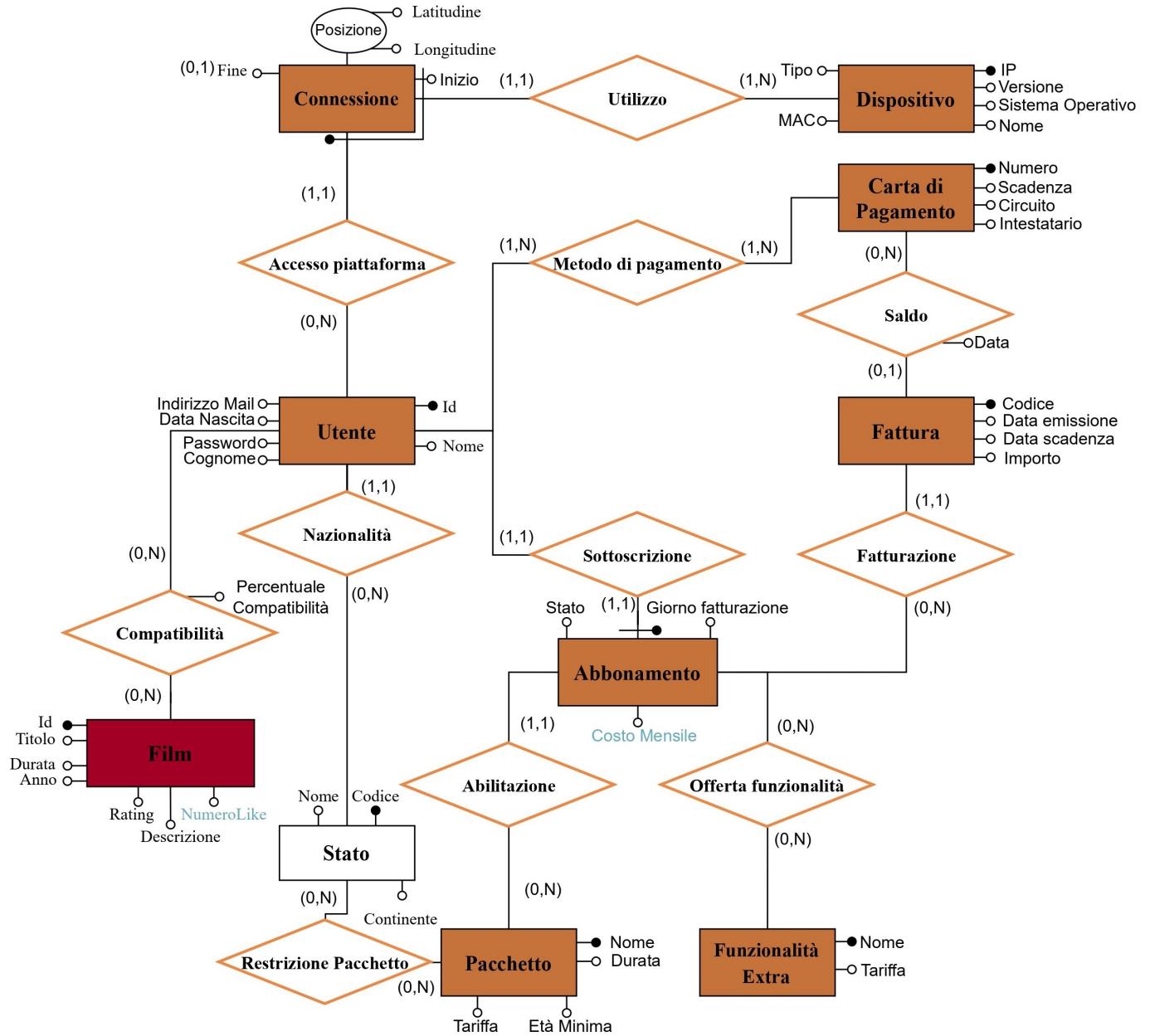
La relationship *Codifica* è stata rappresentata nella suddetta maniera a causa del fatto che un formato video può supportare codec differenti. È possibile avere due contenuti distinti in formato *MP4*, dove uno dei quali utilizza il codec *H.264* per quanto riguarda la codifica dei dati video e il codec *AAC* per quanto riguarda i dati audio, mentre l'altro utilizza il codec *H.265* per quanto riguarda il video e *AC3* per la codifica audio pur avendo lo stesso formato.

3.3.2 Restrizioni Formati

È possibile che un particolare contenuto di un film sia disponibile solamente in determinate zone del mondo. Per quanto riguarda la suddivisione delle località geografiche è stato deciso di utilizzare un livello di granularità basato sugli stati geopolitici, perciò uno Stato è un'area geografica. È dunque possibile andare a inserire una restrizione per un determinato contenuto in un determinato Stato per impedirne la visualizzazione di tale contenuto attraverso la relationship molti a molti *Restrizione Contenuto* tra l'entità *Contenuto* e l'entità *Stato*. Dunque gli utenti che hanno come nazionalità tale Stato non potranno richiedere l'erogazione del contenuto ristretto.

Area Clienti

4.1 Sezione Diagramma Interessato



4.2 Dizionario Dei Dati

Entità	Descrizione	Attributi	Identifieri
Utente	Persona fisica registrata alla piattaforma.	Id, Indirizzo Mail, Cognome , Nome, Data di Nascita, Password	Id
Dispositivo	Device su cui è possibile usufruire del servizio.	Ip, Versione, Sistema Operativo, Nome, MAC, Tipo	Ip
Connessione	Connessione di un utente al servizio attraverso un dispositivo.	Inizio, Fine, Posizione	Id(Utente), Inizio, Ip(Dispositivo)
Carta di Pagamento	Informazioni relative alla carta di pagamento utilizzabile sulla piattaforma.	Numero, Scadenza, Circuito, Intestatario	Numero
Fattura	Informazioni relative alla fattura emessa dalla piattaforma ad un utente.	Codice, Data Emissione, Data Scadenza, Importo	Codice
Abbonamento	Piano di sottoscrizione personalizzato relativo ad un utente.	Stato, Giorno fatturazione, Costo Mensile	Id(Utente)
Pacchetto	Piano di funzionalità offerte dal servizio.	Nome, Tariffa, Età Minima, Durata	Nome
Funzionalità Extra	Funzionalità che può essere aggiunta in un piano di abbonamento.	Nome, Tariffa	Nome

Relationship	Descrizione	Attributi	Entità coinvolte
Nazionalità	Stato di nascita di un utente.		Utente(1,1) Stato(0,N)
Accesso Piattaforma	Accesso di un utente alla piattaforma.		Utente(0,N) Connessione(1,1)
Compatibilità	Affinità di un utente con un film.	Percentuale Compatibilità	Film(0,N) Utente(0,N)

Utilizzo	Connessione tramite un determinato dispositivo.		Connessione(1,1) Dispositivo(1,N)
Metodo di Pagamento	Metodo di pagamento utilizzato da un utente per la fatturazione.		Utente(1,N) Carta di Pagamento(1,N)
Sottoscrizione	Relazione tra un utente e il suo piano di abbonamento personalizzato.		Utente(1,1) Abbonamento(1,1)
Saldo	Informazioni relative al saldo di una fattura.	Data	Carta di Pagamento(0,N) Fattura(0,1)
Fatturazione	Piano di abbonamento associato ad una specifica fattura.		Abbonamento(0,N) Fattura(1,1)
Abilitazione	Abilitazione di un pacchetto in un abbonamento di un utente.		Abbonamento(1,1) Pacchetto(0,N)
Offerta Funzionalità	Funzionalità extra associata ad un abbonamento.		Abbonamento(0,N) Funzionalità Extra(0,N)
Restrizione Pacchetto	Possibile restrizione geografica di un pacchetto.		Stato(0,N) Pacchetto(0,N)

4.3 Descrizione Area

L'area clienti è la zona del database dedicata alla memorizzazione delle informazioni relative agli utenti e alle loro interazioni con la piattaforma.

4.3.1 Clienti e Abbonamenti

Per fornire un'esperienza utente di alta qualità, FilmSphere necessita di immagazzinare i dati relativi a tutti gli utenti come l'*indirizzo email*, *data di nascita*, *password*, *nome* e *cognome*. Ad ogni *Utente* delle piattaforma è associato, in maniera univoca, un *Abbonamento* mensile che presenta uno *stato*, che può essere *attivo* o *interrotto*, un *costo mensile* in euro, e un *giorno di fatturazione*. Dunque la creazione di un'istanza di *Utente* e di un'istanza di *Abbonamento* viene considerata un'operazione atomica. Ciò è dettato dal fatto che tra le due entità è presente una relationship *Sottoscrizione* con cardinalità unaria in entrambe le direzioni. Ad ogni *Abbonamento* è inoltre associato uno dei cinque pacchetti offerti dalla piattaforma e delle funzionalità aggiuntive che possono essere aggiunte

al normale piano di abbonamento. Non deve essere possibile aggiungere una funzionalità extra, se questa è già inclusa nel pacchetto.

In seguito una descrizione delle funzionalità offerte dai pacchetti disponibili.

Pacchetto	FHD	4K	8K	Download	No pubblicità	Erogazioni contemporanee
Basic	✓					1
Premium	✓				✓	2
Pro	✓			✓	✓	4
Deluxe	✓	✓		✓	✓	8
Ultimate	✓	✓	✓	✓	✓	Illimitate

È stata infine considerata la possibilità che un pacchetto non sia reso disponibile in una o più aree geografiche attraverso la relationship molti a molti *Restrizione Pacchetto* tra l'entità *Pacchetto* e l'entità *Stato*.

Dunque gli utenti che hanno come nazionalità uno Stato ristretto per tale pacchetto, o non soddisfano l'età minima richiesta, non potranno abilitare tale piano di abbonamento.

4.3.2 Connessioni e Dispositivi

Per fornire la migliore esperienza di visualizzazione possibile, FilmSphere tiene traccia delle connessioni che vengono instaurate tra l'utente e la piattaforma attraverso l'entità *Connessione*, che memorizza la posizione geografica dalla quale avviene la connessione dell'utente, gli istanti di inizio e di fine e il dispositivo utilizzato per accedere al servizio.

Ogni *Dispositivo* è contraddistinto dal suo indirizzo *Ip*, che viene supposto univoco, e presenta informazioni quali il *sistema operativo* utilizzato e la relativa *versione*, il *nome* del dispositivo, il *tipo*, che può essere *tablet*, *smartphone* o *desktop* e il suo relativo *MAC address*.

La memorizzazione delle connessioni e dei dispositivi utilizzati da ciascun utente è cruciale per comprendere le abitudini di visualizzazione degli utenti e quali sono i loro dispositivi preferiti.

4.3.3 Fatturazioni e Pagamenti

Per tenere traccia dei pagamenti e permettere agli utenti di saldare le loro fatture è necessario consentire all'utente di inserire molteplici istanze di *Carta di Pagamento*, attraverso le quali sarà in grado di saldare le fatture relative all'abbonamento in corso. L'entità *Carta di Pagamento* memorizza il *numero* della carta, la data di *scadenza*, il relativo *circuito* tra quelli ammessi dalla piattaforma, e l'*intestatario*, che può differire dal proprietario dell'abbonamento stesso. La memorizzazione delle fatture è resa possibile dall'implementazione

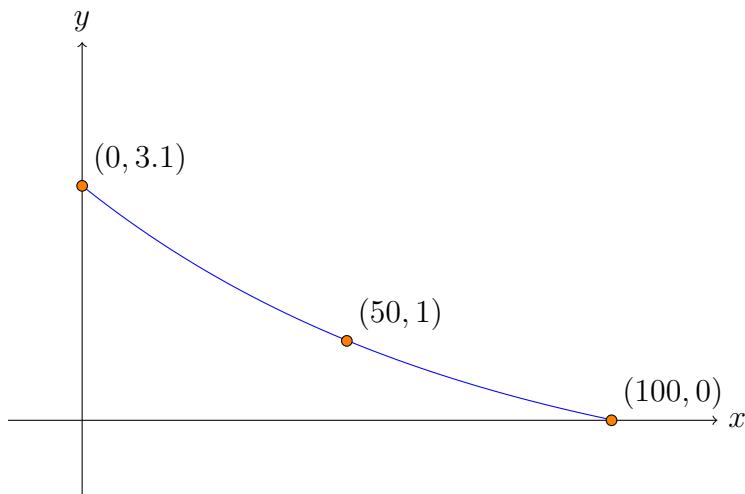
di un'entità *Fattura* che contiene la *data emissione* e la *data scadenza* della fattura, il relativo *importo* e l'abbonamento a cui si riferisce.

Una volta che una fattura viene saldata, viene aggiunta un'istanza della relationship *Saldo* che indica che una tale *Fattura* è stata saldata utilizzando una *Carta di Pagamento* in una determinata *data*, e questo sarà possibile solo se la carta utilizzata è un'istanza della relationship *Metodo di Pagamento* relativa all'utente intestatario della fattura.

4.3.4 Sistema di Raccomandazione Contenuti

FilmSphere ha optato per l'implementazione di un sistema di raccomandazioni personalizzate basate sulla percentuale di compatibilità tra un determinato film e un utente. Dunque ogni utente registrato alla piattaforma avrà una determinata percentuale di compatibilità con tutti i film della piattaforma, in modo da essere affiancato durante il suo processo di scelta del contenuto da visionare. Le percentuali di compatibilità di un determinato utente vengono aggiornate ogni volta che viene terminata la visualizzazione di un contenuto, o viene lasciata una recensione ad un contenuto video.

La percentuale di compatibilità di default per un utente che si registra alla piattaforma è del 50% per ogni contenuto video. Per effettuare gli incrementi e i decrementi FilmSphere utilizza il moltiplicatore λ che permette di uniformare questi ultimi alla percentuale di compatibilità precedente ed è definito dalla seguente formula.



$$\lambda(x) = 4.2 \cdot e^{(-\frac{1}{75}x)} - 1.1$$

Dove :

- In caso di incremento $x = \text{Percentuale Precedente}$
- In caso di decremento $x = 100 - \text{Percentuale Precedente}$

Ogni incremento e decremento percentuale descritto successivamente dovrà essere calcolato attraverso il prodotto tra il moltiplicatore λ , che ha come input la percentuale di compatibilità precedente, e l'incremento o decremento stesso.

Il valore di λ varia tra 0 e 3.1, quando il valore di input x varia tra 0 e 100, mentre quando si ha una percentuale di compatibilità di default (50%) λ ha valore 1.

Questa funzione infatti permette, grazie all'avvicinamento esponenziale all'asse y , di ammortizzare gli incrementi delle compatibilità elevate e i decrementi di quelle irrisorie, e di accentuare i decrementi delle compatibilità elevate e gli incrementi di quelle più basse.

Visualizzazione Terminata

Quando viene terminata la visualizzazione di un contenuto ne viene valutata la percentuale di visualizzazione. In base a tale valore la percentuale di compatibilità dei contenuti dello stesso genere viene impattata secondo la seguente logica:

- Se la percentuale è inferiore al 5%: le compatibilità dei film dello stesso genere rimangono invariate.
- Se la percentuale è compresa tra il 5% e il 60%:

Le compatibilità dei film dello stesso genere subiscono un decremeento di:
 $[(100 - \text{Percentuale visualizzazione})/10]\%$

Il film visualizzato subisce un ulteriore decremeento di: 5%

- Se la percentuale è superiore al 60%:

Le compatibilità dei film dello stesso genere subiscono un incremento di:
 $[(\text{Percentuale visualizzazione})/10]\%$

Il film visualizzato subisce un ulteriore incremento di: 5%

I 5 migliori film dello stesso genere in base al rating subiscono un ulteriore incremento di: 6%

Nuova Recensione Aggiunta

Quando viene aggiunta una nuova recensione da parte di un utente ad un determinato contenuto, i film dello stesso genere e quelli diretti dallo stesso regista subiranno un incremento o decremeento a seconda dell'esito della recensione.

- In caso di recensione positiva:

Tutti i film dello stesso genere subiscono un incremento del 5% di compatibilità.

I film dello stesso regista subiscono un ulteriore incremento del 3% di compatibilità.

Il film stesso subisce un incremento del 10% di compatibilità.

- In caso di recensione negativa:

Tutti i film dello stesso genere subiscono un decremeento del 5% di compatibilità.

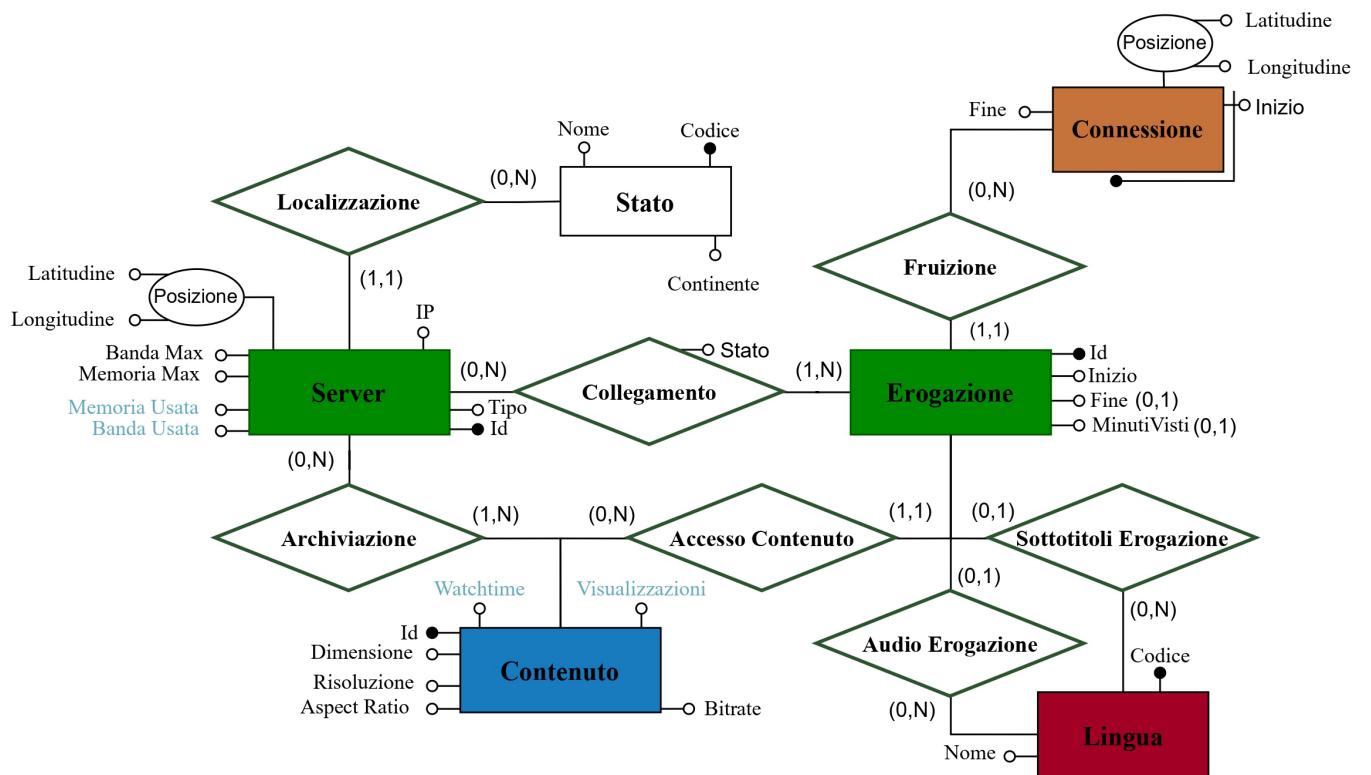
I film dello stesso regista subiscono un ulteriore decremeento del 3% di compatibilità

Il film stesso subisce un decremeento del 10% di compatibilità.

A livello implementativo, le percentuali di compatibilità personalizzate per ogni utente sono gestite attraverso la relationship molti a molti *Compatibilità* tra le entità *Film* e l'entità *Utente*, che presenta come attributo la *percentuale di compatibilità* tra l'utente e il film stesso.

Area Streaming

5.1 Sezione Diagramma Interessato



5.2 Dizionario Dei Dati

Entità	Descrizione	Attributi	Identifieri
Server	Macchina fisica connessa in rete in grado di fornire contenuti attraverso flussi streaming.	Id, Banda Max, Memoria Max, Memoria Usata, Banda Usata, Tipo, Ip, Posizione	Id
Erogazione	Visualizzazione di un contenuto attraverso un server da parte di un utente connesso alla piattaforma.	Id, Inizio, Fine, Minuti visti	Id

Relationship	Descrizione	Attributi	Entità coinvolte
Accesso Contenuto	Accesso ad un contenuto video.		Contenuto(0,N) Erogazione(1,1)
Collegamento	Erogazione da parte di un server.	Stato	Erogazione(1,N) Server(0,N)
Fruizione	Fruizione di un'erogazione da una connessione.		Connessione(0,N) Erogazione(1,1)
Archiviazione	Possesso in memoria di un contenuto da parte di un server.		Contenuto(1,N) Server(0,N)
Localizzazione	Geolocalizzazione di un server in uno Stato.		Server(1,1) Stato(0,N)
Audio Erogazione	Audio con cui è stato visualizzato un determinato contenuto.		Erogazione(0,1) Lingua(0,N)
Sottotitoli Erogazione	Sottotitoli con cui è stato visualizzato un determinato contenuto.		Erogazione(0,1) Lingua(0,N)

5.3 Descrizione Area

L'area formati è la zona del database dedicata alla memorizzazione delle informazioni relative alla Content Delivery Network.

5.3.1 Server e Memorizzazione

Per fornire un servizio di streaming senza compromessi FilmSphere ha a disposizione una rete di distribuzione contenuti sparsa in tutto il mondo che permette agli utenti di usufruire dei contenuti disponibili sulla piattaforma senza rallentamenti o malfunzionamenti.

La CDN di FilmSphere dovrà avere a disposizione un quantitativo di server sparsi per il globo e posizionati in base alla quantità e alla distribuzione del pubblico della piattaforma, in modo da gestire al meglio le richieste di contenuto. Un *Server* è contraddistinto da un *identificatore* univoco, una *posizione* geografica espressa in coordinate GPS, la tipologia che può essere *Edge Server* o *Origin Server*, una *banda massima* fornita dal provider espressa in Mbit/s, una *memoria massima* di archiviazione per contenuti espressa in GB e un indirizzo *IP*.

Ogni server, come indicato precedentemente, è localizzato fisicamente in una particolare zona geografica. È stato scelto di implementare sia la località geografica in coordinate GPS, attraverso un attributo, che lo Stato in cui il server si trova, attraverso la relationship *Localizzazione*, in modo da poter sia effettuare delle operazioni che riguardano la nazione dove risiede il server, sia la possibilità di conoscere l'effettiva posizione geografica per gestire al meglio i collegamenti tra gli utenti e il server stesso.

Ogni server è in grado di immagazzinare dei contenuti multimediali, entro la relativa *memoria massima*, e può essere responsabile di un'erogazione di un determinato contenuto. Il possesso da parte di un server di un contenuto è espresso attraverso la relationship molti a molti *Archiviazione* tra l'entità *Contenuto* e l'entità *Server*.

Non deve essere possibile eliminare un'istanza di *Archiviazione* relativa ad un server, se il contenuto archiviato è attualmente erogato dal server stesso.

5.3.2 Fruizione ed Erogazione

Un'*Erogazione* rappresenta una visualizzazione di un contenuto, attraverso un *Collegamento* ad un *Server* da parte di un *Utente*, identificato attraverso una *Connessione* attiva al servizio. Un'*Erogazione* è contraddistinta da un timestamp di *inizio* e di *fine* che permettono al sistema di valutare i tempi d'erogazione medi di un determinato contenuto.

Inoltre l'entità *Erogazione* presenta l'attributo *minuti visti* che permette di valutare l'effettiva durata della visualizzazione del contenuto, a causa delle possibili pause durante l'erogazione del contenuto stesso.

Le relationship *Audio Erogazione* e *Sottotitoli Erogazione* permettono di identificare quelle che sono la traccia audio e i sottotitoli utilizzati durante la fruizione del contenuto. Entrambe le relationship presentano cardinalità opzionale unaria per quanto riguarda l'entità *Erogazione* a causa del fatto che un film potrebbe non avere delle lingue disponibili, dunque essere muto, e potrebbe non essere richiesto l'utilizzo di sottotitoli da parte dell'utente. Occorre inoltre controllare che le lingue selezionate per la traccia audio e i sottotitoli siano disponibili per il film in questione, e che dunque siano istanze della relationship *Audio* e *Sottotitoli* per quanto riguarda il film relativo al contenuto.

La relationship *Collegamento* ha cardinalità uno a molti per quanto riguarda l’entità *Erogazione* a causa del fatto che durante un’erogazione di un contenuto è possibile che avvenga un cambio di server in seguito ad un possibile sovraccarico o malfunzionamento.

Inoltre la relationship *Collegamento* possiede un attributo *stato* che può essere *attivo*, *interrotto* o *terminato*, che permette di identificare in un dato momento dell’erogazione qual’è il server utilizzato e, eventualmente, quelli utilizzati in precedenza. Una volta che l’erogazione è terminata il collegamento viene automaticamente considerato terminato. Un’erogazione può avere al massimo un collegamento attivo alla volta, e se questo viene interrotto non sarà più possibile riattivarlo.

Immagazzinare i dati relativi alle erogazioni permette di andare a implementare il sistema di raccomandazioni personalizzato e tenere traccia di tutte le visualizzazioni dei contenuti presenti sulla piattaforma.

5.3.3 Caching

Il caching è una tecnica che viene utilizzata per migliorare l’efficienza e la velocità di accesso ai contenuti da parte degli utenti. FilmSphere ha implementato una strategia di caching che permette di andare a gestire la memoria di archiviazione dei server andando a prevedere la possibile domanda futura di un contenuto multimediale.

L’operazione di gestione della cache propone l’archiviazione dei contenuti relativi ai tre film con la compatibilità maggiore per ciascun utente, indicando per ciascun contenuto il server su cui fare il caching del contenuto. Il server più propenso all’erogazione di quei contenuti è quello che presenta la distanza minore dalla posizione dell’ultima connessione alla piattaforma di ciascun utente. Di conseguenza sarà tale server a ricevere tutti i contenuti relativi ai film con compatibilità maggiore, solo se non già presenti nella cache del server.

Grazie a questa funzionalità sarà possibile spostare in anticipo i contenuti ed evitare rallentamenti o sovraccarichi dei server della CDN durante la fase di utilizzo.

5.3.4 Richiesta Contenuti e Scelta del Server

Per fare in modo che la CDN sia in grado di gestire un elevato quantitativo di flussi di streaming simultanei è necessario andare a implementare una strategia di bilanciamento del carico, che permetta di selezionare il server più opportuno per ogni utente.

Il flow di richiesta di un contenuto multimediale è definito dai seguenti passaggi:

1. L’utente effettua il login alla piattaforma, stabilendo una connessione
2. L’utente richiede la fruizione di un determinato contenuto relativo ad un film
3. Si ricerca il server che possiede tale contenuto, è in grado di effettuare lo streaming e ha il costo di streaming minore
4. Viene creata l’erogazione del contenuto relativa al server e alla connessione dell’utente

È dunque necessario definire una metrica che definisca il costo relativo allo streaming di un determinato contenuto tra un server ed un utente.

Il carico di un server è definito come la percentuale di banda pubblica utilizzata in quell'esatto momento. Si assume che i server della CDN abbiano a disposizione un hardware che permetta di sfruttare al massimo la capacità di streaming fornita dal provider.

$$\text{Carico Server} = \frac{\text{Used Bandwidth}}{\text{Max Public Bandwidth}}$$

Costo di Streaming

Il costo di streaming di un contenuto multimediale è definito in base al carico del server, alla banda necessaria per effettuare lo streaming e alla distanza tra il server e l'utente finale.

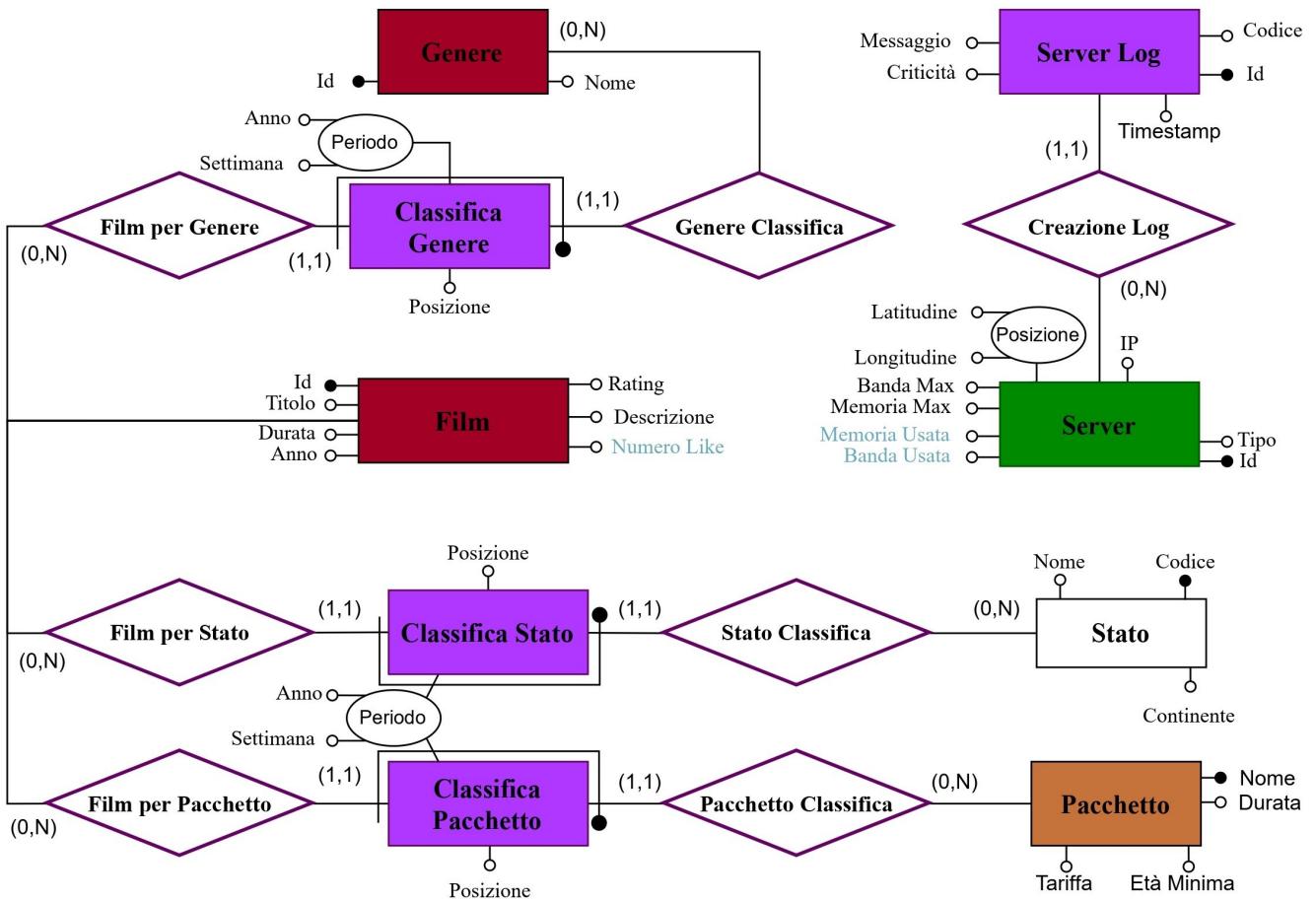
- Carico del Server. **Peso:** 50%.
- Banda necessaria per effettuare lo streaming rispetto alla banda libera:
Bitrate/(Max Public Bandwidth – Used Bandwidth). **Peso:** 20%.
- Distanza dall'utente: Distanza Utente-Server/Distanza Massima. **Peso:** 30%.

$$\text{Costo di Streaming} = \sum \text{Fattore} \cdot \text{Peso}$$

Dunque, per andare a valutare quale server è il più indicato per andare ad effettuare lo streaming, occorre andare a valutare il costo di streaming per ognuno di essi e scegliere di conseguenza quello con costo minore.

Area Analytics

6.1 Sezione Diagramma Interessato



6.2 Dizionario Dei Dati

Entità	Descrizione	Attributi	Identifieri
Server Log	Log emesso da un server della CDN.	Id, Codice, Criticità, Messaggio, Timestamp	Id
Classifica Genere	Classifica relativa ai diversi generi cinematografici.	Periodo, Posizione	Id(Film), Id(Genere), Periodo
Classifica Stato	Classifica relativa agli Stati.	Periodo, Posizione, Risoluzione	Id(Film), Codice(Stato), Periodo
Classifica Pacchetto	Classifica relativa ai pacchetti disponibili.	Periodo, Posizione, Risoluzione	Id(Film), Nome(Pacchetto), Periodo

Relationship	Descrizione	Attributi	Entità coinvolte
Creazione Log	Relazione tra un log e un server.		Server(0,N) Server Log(1,1)
Film per Stato	Relazione tra Classifica Stato e Film.		Film(0,N) Classifica Stato(1,1)
Film per Pacchetto	Relazione tra Classifica Pacchetto e Film.		Film(0,N) Classifica Pacchetto(1,1)
Film per Genere	Relazione tra Classifica Genere e Film.		Film(0,N) Classifica Genere(1,1)
Stato Classifica	Relazione tra Classifica Stato e Stato.		Stato(0,N) Classifica Stato(1,1)
Pacchetto Classifica	Relazione tra Classifica Pacchetto e Pacchetto.		Pacchetto(0,N) Classifica Pacchetto(1,1)
Genere Classifica	Relazione tra Classifica Genere e Genere.		Genere(0,N) Classifica Genere(1,1)

6.3 Descrizione Area

L'area analytics è la zona del database dedicata alla memorizzazione di informazioni utili al miglioramento dell'esperienza utente e dell'efficienza della CDN.

FilmSphere ha ritenuto opportuno, in seguito alla presenza di molteplici funzionalità analytics lato server, necessarie per il funzionamento della piattaforma, l'implementazione di una sezione del database volta alla memorizzazione di tali risultati e informazioni utili, cosicché possano essere fruite in maniera differita rispetto alla loro valutazione.

In questa sezione verranno discusse solamente le implementazioni dal punto di vista concettuale e non si terrà conto della reale implementazione delle funzionalità stesse, che verranno ampiamente discusse nella sezione dedicata (12).

6.3.1 Classifiche

FilmSphere ha un'area dedicata alla memorizzazione di classifiche aggiornate settimanalmente dei contenuti per migliorare l'esperienza utente. Le classifiche in questione sono 3 e sono rappresentate dalle entità: *Classifica Genere*, *Classifica Pacchetto*, *Classifica Stato*.

Ogni entità in questione presenta un attributo composto *periodo*, che permette di identificare la settimana e l'anno in cui tale classifica è stata valutata e la *posizione* relativa in quel determinato periodo.

Per quanto riguarda *Classifica Stato* e *Classifica Pacchetto* si ha anche un attributo *risoluzione* che indica la risoluzione del contenuto più richiesto per tale film.

6.3.2 Bilanciamento del carico

Per lo sviluppo della funzionalità di bilanciamento del carico, il database è provvisto di una funzionalità di immagazzinamento dei log forniti dai server della CDN durante il loro funzionamento.

Tali log sono memorizzati attraverso l'entità *Server Log* che presenta un *timestamp* relativo al momento di erogazione del log, un *messaggio*, un *codice*, e il livello di *criticità* del log stesso, che può essere *info*, *warning* oppure *error*.

I log in questione permetteranno di analizzare approfonditamente lo stato dei server durante il funzionamento del servizio e di conseguenza saranno fondamentali per stimare dei possibili sovraccarichi futuri e gestire il bilanciamento del carico dell'intera CDN.

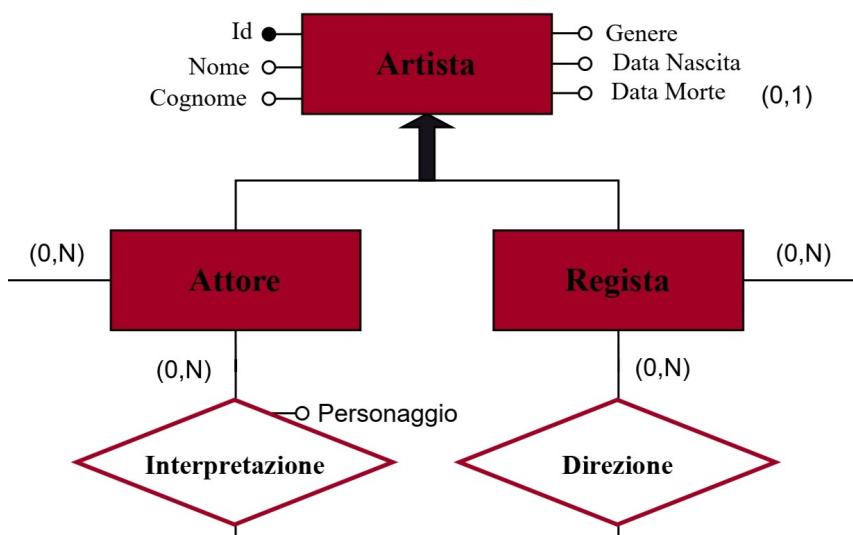
Parte II

Ristrutturazione

Eliminazione Generalizzazioni

7.1 Generalizzazione di Artista

La generalizzazione totale *Artista* rappresenta un individuo facente parte del cast o della regia, che è dunque coinvolto in prima persona nella realizzazione di un film. La generalizzazione viene considerata totale a seguito del fatto che la realtà d'analisi tiene solamente in considerazione i membri del cast e della regia, e non l'intera troupe cinematografica che partecipa alla produzione del film. Di conseguenza un artista deve essere necessariamente un attore o un regista.

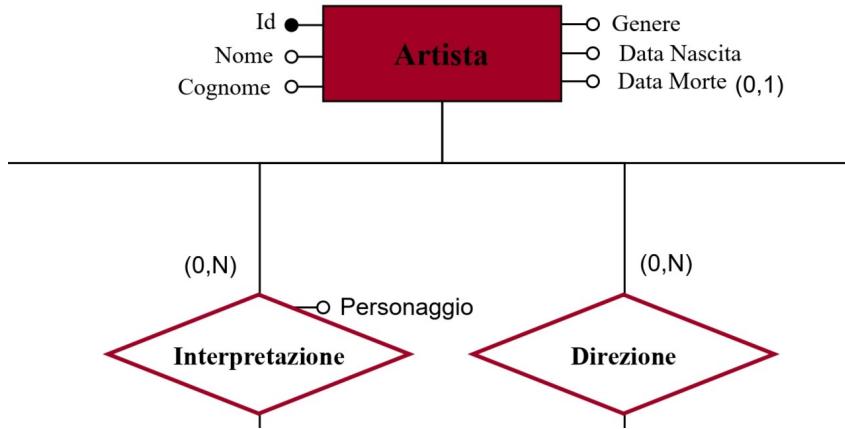


È stato optato per una ristrutturazione della generalizzazione attraverso l'accorpamento delle entità figlie *Attore* e *Regista* nell'entità genitrice *Artista*.

La motivazione principale che ha portato a tale decisione è la sovrapposizione della generalizzazione: circa il 30% degli artisti hanno lavorato nel corso della loro carriera sia come artisti, che come registi. Dunque questo porterebbe ad una doppia memorizzazione di una buona parte degli artisti, con un conseguente spreco di memoria. Inoltre l'assenza di attributi discordanti tra le entità figlie rende la scelta ancor più evidente.

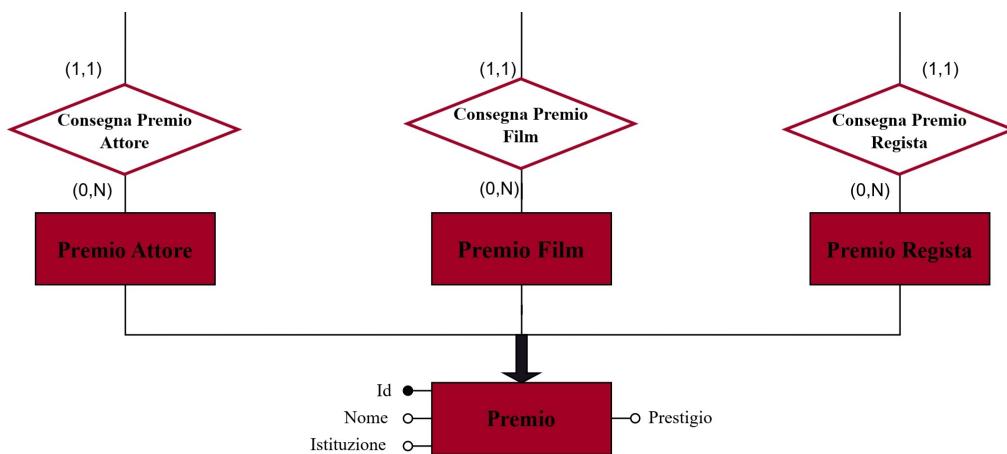
Dunque il risultato di questa operazione di ristrutturazione è la modifica delle relationship *Interpretazione*, *Direzione*, *Nomina Attore* e *Nomina Regista* che non comprenderanno più le entità figlie *Attore* e *Regista* ma saranno direttamente collegate con l'entità genitrice *Artista*.

Questa soluzione non provoca nessun tipo di perdita di informazione perché le relationship sopra citate permettono in egual maniera di identificare il ruolo di un determinato artista nella produzione di un dato film.



7.2 Generalizzazione di Premio

La generalizzazione totale *Premio* rappresenta una premiazione attribuita ad un regista, attore o film stesso da parte di un'istituzione. La generalizzazione viene considerata totale a seguito del fatto che un premio registrato sulla piattaforma deve appartenere necessariamente ad una di queste tre categorie. Inoltre la generalizzazione è esclusiva, dato che un singolo premio può appartenere solamente ad una delle tre categorie sopra citate.



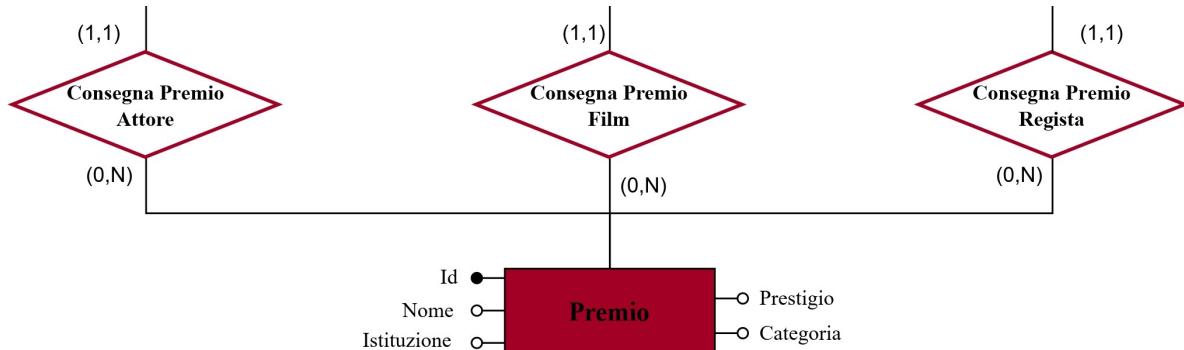
In seguito a tali considerazioni è stato optato per una ristrutturazione della generalizzazione attraverso l'accorpamento delle entità figlie *Premio Attore*, *Premio Regista* e *Premio Film* nell'entità genitrice *Premio*.

Per mantenere l'integrità del database, è stato ritenuto necessario introdurre l'attributo *categoria* all'interno dell'entità *Premio*. Questo permette al sistema, ogni qualvolta venga

creata un'istanza di una relationship riguardante una premiazione, di controllare se il premio attribuito rispecchia le caratteristiche della relationship stessa.

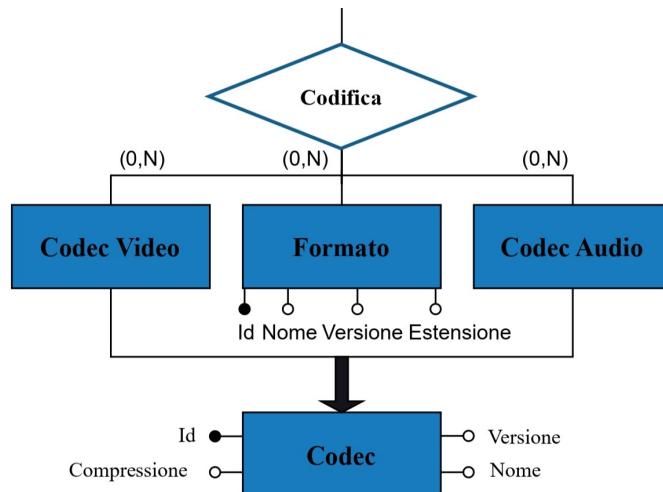
Inoltre, per quanto riguarda la creazione di un'istanza delle relationship *Premio Attore* e *Premio Regista* è necessario controllare precedentemente se tale artista ha realmente preso parte alla realizzazione del film con tale ruolo.

Dunque il risultato della ristrutturazione consiste nel coinvolgimento diretto delle relationship *Consegna Premio Attore*, *Consegna Premio Regista* e *Consegna Premio Film* con l'entità genitrice *Premio*.



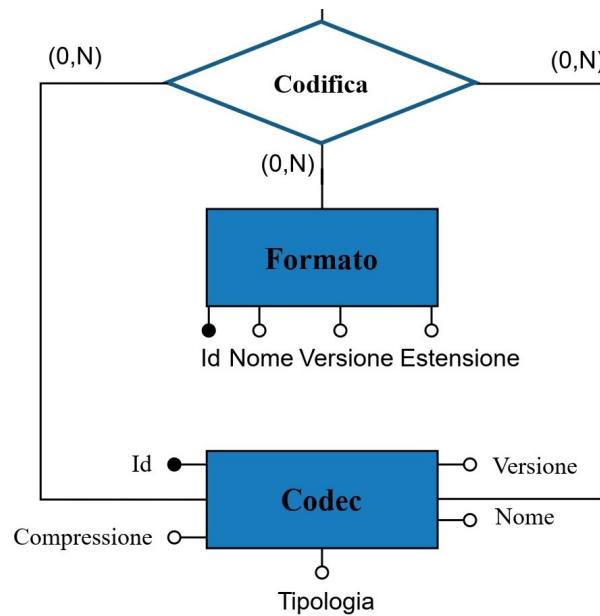
7.3 Generalizzazione di Codec

La generalizzazione totale di *Codec* rappresenta un tipologia di Codifica per quanto riguarda un contenuto multimediale. La gestione delle ristrutturazione della generalizzazione prende in analisi le stesse considerazioni effettuate per quanto riguarda la generalizzazione dell'entità *Premio* (7.2). Allo stesso modo, la generalizzazione è totale e esclusiva perché un singolo codec deve riguardare necessariamente la codifica audio, oppure la codifica video.



Anche in questo caso è stato ritenuto necessario introdurre un nuovo attributo all'entità *Codec* chiamato *tipologia* per differenziare le due diverse tipologie di Codec *audio* e *video*.

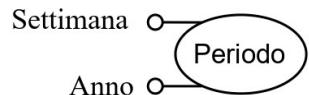
Per quanto riguarda la creazione di un'istanza della relationship *Codifica* è necessario controllare che comprenda un'istanza di Codec di tipo audio e un'istanza di Codec di tipo video.



Eliminazione Attributi Composti

8.1 Attributo composto Periodo

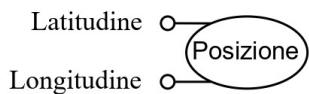
È stato realizzato un attributo composto *periodo* per essere in grado di memorizzare la coppia di attributi *settimana* e *anno* relativi ad una determinata istanza delle entità *Classifica Genere*, *Classifica Stato* e *Classifica Pacchetto*.



È stato deciso di tradurre l'attributo composto inserendo nelle entità tutti gli attributi di *periodo*, in quanto non è stato ritenuto necessario trasformarlo in un'entità a causa del fatto che ogni volta che si accede ad una classifica è necessario conoscere il periodo alla quale si riferisce.

8.2 Attributo composto Posizione

È stato realizzato un attributo composto *posizione* per essere in grado di memorizzare la coppia di attributi *latitudine* e *longitudine* relativi ad una determinata istanza delle entità *Server* e *Connessione*.



Per la traduzione si è optato per la stessa strategia utilizzata per l'attributo *periodo*, andando ad inserire nelle entità tutti gli attributi di *posizione*, in quanto ogni volta che si accede ad una server o ad una connessione è necessario avere informazioni sulla loro geolocalizzazione.

Parte III

Analisi Prestazionale

Tavola dei Volumi

Concetto	Tipo	Volume	Considerazioni
Area Generale			
Stato	E	196	Il servizio sarà disponibile in 196 Paesi.
Area Contenuti			
Film	E	400	Ipotizzato.
Artista	E	515	Ipotizzando 480 attori e 50 registi, e ipotizzando che il 30% dei registi siano anche attori.
Genere	E	20	Ipotizzato.
Lingua	E	20	Ipotizzato.
Critico	E	50	Ipotizzato.
Premio	E	60	Ipotizzando che ogni anno vengano assegnati 20 premi per i film, 30 per gli attori e 10 per i registi.
Premiazione Film	E	30	Ipotizzando che il 50% dei film premiati ogni anno appartenga alla piattaforma, e che siano archiviati i dati degli ultimi 3 anni.
Premiazione Attore	E	45	Ipotizzando che il 50% dei premi assegnati agli attori ogni anno sia per film appartenenti alla piattaforma, e che siano archiviati i dati degli ultimi 3 anni.
Premiazione Regista	E	15	Ipotizzando che il 50% dei registi premiati ogni anno sia per film appartenenti alla piattaforma, e che siano archiviati i dati degli ultimi 3 anni.
Consegna Premio Film	R	30	Stesso volume di Premiazione Film.
Consegna Premio Attore	R	45	Stesso volume di Premiazione Attore.
Consegna Premio Regista	R	15	Stesso volume di Premiazione Regista.
Film Interpretato	R	45	Stesso volume di Premiazione Attore.
Film Diretto	R	15	Stesso volume di Premiazione Regista.
Nomina Attore	R	45	Stesso volume di Premiazione Attore.
Nomina Film	R	30	Stesso volume di Premiazione Film.

Nomina Regista	R	15	Stesso volume di Premiazione Regista.
Interpretazione	R	6.000	Ipotizzando che un attore reciti in media in 15 film diversi. $400 * 15 = 6.000$.
Direzione	R	500	Ipotizzando che in media un film ogni 4 abbia oltre che il regista principale anche un regista secondario.
Classificazione	R	800	Ipotizzando che un film in media appartenga a 2 generi diversi.
Recensione Critico	R	14.000	Ipotizzando che un film in media venga recensito dal 70% dei critici della piattaforma, dunque per ogni film ci saranno 35 recensioni di critici. $35 * 400 \text{ film} = 14.000$.
Recensione Utente	R	2.160.000	Ipotizzando che su 10.000 visualizzazioni giornaliere, circa il 20% di queste siano accompagnate anche da una recensione. $10.000 * 0.2 = 2.000$ recensioni al giorno, 60.000 al mese, 720.000 all'anno. Ipotizzando i dati degli ultimi 3 anni saranno presenti nel database circa $720.000 * 3 \text{ anni} = 2.160.000$ recensioni.
Sottotitoli	R	6.000	Ipotizzando che un film in media abbia i sottotitoli disponibili in 15 lingue, $15 * 400 = 6.000$.
Audio	R	3.200	Ipotizzando che un film sia disponibile in media in 8 lingue, $8 * 400 = 3.200$.
Produzione	R	400	Stesso volume di Film.
Area Clienti			
Utente	E	10.000	Ipotizzato.
Abbonamento	E	10.000	Stesso volume di Utente.
Pacchetto	E	5	La piattaforma supporta 5 modalità di abbonamento.
Funzionalità Extra	E	10	Ipotizzato.
Carta di Pagamento	E	10.900	Ipotizzando che il 10% degli utenti possiede 2 carte registrate e che l'1% degli utenti abbiano registrato una carta di pagamento già registrata da un altro utente.
Fattura	E	360.000	Ipotizzando solo le fatture degli ultimi 3 anni e considerando una fatturazione mensile si hanno 10.000 abbonamenti mensili attivi $* 12 \text{ mesi} * 3 \text{ anni} = 360.000$.

Connessione	E	10.800.000	Ipotizzando che in media un utente si connette alla piattaforma una volta al giorno. Considerando le connessioni degli ultimi 3 anni si hanno: 10.000 utenti * 30 giorni = 300.000 connessioni al mese, 300.000 * 12 mesi = 3.600.000 all'anno, 3.600.000 * 3 anni = 10.800.000 connessioni totali.
Dispositivo	E	20.000	In media un utente si connette alla piattaforma utilizzando 2 diversi dispositivi.
Metodo di Pagamento	R	11.000	Considerando che il 10% degli utenti possiede 2 carte registrate.
Nazionalità	R	10.000	Stesso volume di Utente.
Saldo	R	350.000	Considerando come già saldate tutte le fatture fino all'ultimo mese e non saldate le fatture dell'ultimo mese.
Accesso piattaforma	R	10.800.000	Stesso volume di Connessione.
Utilizzo	R	10.800.000	Stesso volume di Connessione.
Fatturazione	R	360.000	Stesso volume di Fattura.
Sottoscrizione	R	10.000	Stesso volume di Utente.
Offerta funzionalità	R	20.000	Ipotizzando 2 funzionalità aggiuntive inserite in media in ciascun abbonamento.
Abilitazione	R	10.000	Stesso volume di Abbonamento.
Restrizione Pacchetto	R	50	Ipotizzando che un pacchetto non è fruibile in media in 10 Stati.
Compatibilità	R	2.000.000	Ipotizzando che un utente abbia una percentuale di compatibilità diversa da quelle di default per la metà dei film sulla piattaforma. 200 film * 10.000 utenti = 2.000.000.
Area Formati			
Contenuto	E	2.000	Ipotizzando che ogni film in media abbia 5 contenuti diversi che lo rappresentano, 5 * 400 = 2.000 contenuti.
Codec	E	10	Ipotizzando 5 diversi codec audio e 5 diversi codec video disponibili.
Formato	E	5	Ipotizzato.
Codifica	R	2.000	Stesso volume di Contenuto.
Restrizione Contenuto	R	20.000	Ipotizzando che un contenuto in media non sia fruibile in 10 Paesi, 10 * 2.000 = 20.000.
Rappresentazione	R	2.000	Stesso volume di Contenuto.
Area Streaming			
Server	E	30	Ipotizzato 1*.

Erogazione	E	10.800.000	Considerando la precedente ipotesi sulle connessioni giornaliere medie di 10.000, ipotizziamo che ad ogni connessione si richieda l'erogazione di un contenuto. Considerando le erogazioni degli ultimi 3 anni saranno lo stesso volume delle connessioni.
Localizzazione	R	30	Stesso volume di Server.
Archiviazione	R	7.500	Ipotizzando che ogni server contenga in media 250 contenuti, $250 * 30 = 7.500$.
Accesso Contenuto	R	10.800.000	Stesso volume di Erogazione.
Collegamento	R	11.016.000	Ipotizzando che per ogni erogazione ci sia un collegamento e che ogni 50 erogazioni ci sia un cambio di server. $10.800.000 + 10.800.000 / 50 = 11.016.000$.
Fruizione	R	10.800.000	Stesso volume di Erogazione.
Audio Erogazione	R	9.720.000	Ipotizzando che su tutte le erogazioni il 90% siano state visualizzate con un audio, mentre il restante 10% siano film muti.
Sottotitoli Erogazione	R	4.320.000	Ipotizzando che su tutte le erogazioni il 40% siano state visualizzate con dei sottotitoli.
Area Analytics			
Server Log	E	972.000	Ipotizzando che ogni server generi 30 log giornalieri, per 30 server saranno 900 Log al giorno, 27.000 al mese, 324.000 l'anno. Considerano l'archivio di 3 anni di piattaforma ci saranno 972.000 log.
Classifica Stato	E	3.120.000	Considerando classifiche settimanali, in un anno ci sono 52 settimane. Ipotizzando che gli utenti realmente registrati appartengano realmente a circa 100 stati diversi. Per ogni Stato in media i film visualizzati si ipotizza siano 200, perciò ogni classifica settimanale sarà formata da $200 \text{ film} * 100 \text{ stati} = 20.000 \text{ record}$, $20.000 * 52 \text{ settimane} * 3 \text{ anni} = 3.120.000$.

Classifica Pacchetto	E	234.000	Avendo 5 pacchetti, viene stilata la classifica dei film più visti da utenti di ciascun pacchetto settimanalmente. Per ogni pacchetto, in media i film visualizzati si ipotizzano 300, perciò ogni classifica settimanale sarà formata da 300 film * 5 pacchetti = 1.500 record, 1.500 * 52 settimane * 3 anni = 234.000.
Classifica Genere	E	15.600	Avendo 20 generi, viene stilata la classifica dei 5 film con rating più alto per ciascun genere settimanalmente. 5 * 20 generi = 100 record alla settimana. 100 * 52 settimane * 3 anni = 15.600.
Creazione Log	R	972.000	Stesso volume di Server Log.
Film per Stato	R	3.120.000	Stesso volume di Classifica Stato.
Film per Pacchetto	R	234.000	Stesso volume di Classifica Pacchetto.
Film per Genere	R	15.600	Stesso volume di Classifica Genere.
Stato Classifica	R	3.120.000	Stesso volume di Classifica Stato.
Pacchetto Classifica	R	234.000	Stesso volume di Classifica Pacchetto.
Genere Classifica	R	15.600	Stesso volume di Classifica Genere.

1* Valutazione volume entità Server:

Ipotizzando che un singolo server abbia una banda pubblica di 10 Gbit/sec e il bitrate medio di un contenuto sulla piattaforma sia di 20 Mbit/sec. Il peso di ogni contenuto è in media 20 GB, quindi in totale nella piattaforma ci saranno 40 TeraByte di contenuti.

In media un server avrà quindi 10 TeraByte di memoria disponibile e sarà in grado di contenere e potenzialmente, grazie alla banda pubblica, streammare circa 500 contenuti contemporaneamente.

Considerando che in media 10.000 contenuti vengono streammati contemporaneamente al giorno, il quantitativo minimo di server per garantire il servizio è di 20 server. È stata dunque raggiunta la conclusione che 30 server possano essere un numero ragionevole per la quantità di utenti considerata.

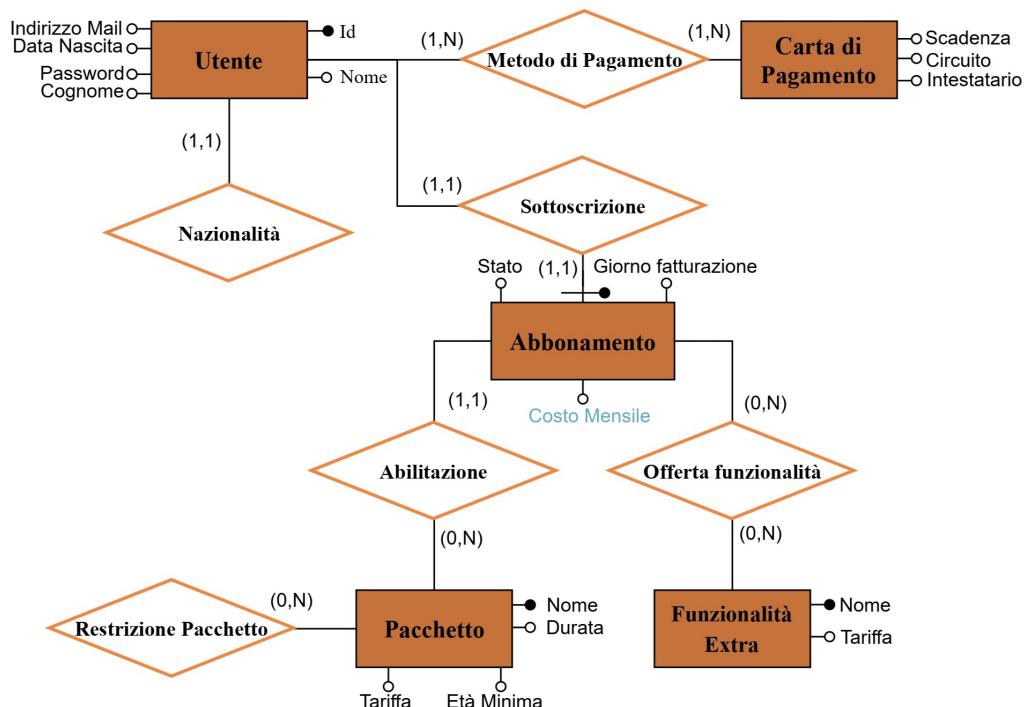
Operazioni sui Dati

10.1 Registrazione Utente e Creazione Abbonamento

Come prima operazione significativa è stata scelta la registrazione di un nuovo utente con la contestuale creazione dell'abbonamento e dell'inserimento del metodo di pagamento necessario per la fatturazione. L'operazione permette dunque di andare a creare un nuovo utente sulla piattaforma e di associarvi un abbonamento relativo ad uno dei pacchetti disponibili, andando inserire una carta di pagamento, e controllando che non vi siano né restrizioni d'età, né restrizioni dal punto di vista geografico.

Input	Output	Frequenza
Attributi di Utente , Attributi di Carta di Pagamento , Nome (Pacchetto), Nome (Funzionalità Extra)	idUtenteRegistrato	50 al giorno

10.1.1 Sezione Diagramma Interessato



10.1.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Clienti			
Utente	E	10.000	Ipotizzato.
Abbonamento	E	10.000	Stesso volume di Utente.
Pacchetto	E	5	La piattaforma supporta 5 modalità di abbonamento.
Funzionalità Extra	E	10	Ipotizzato.
Carta di Pagamento	E	10.900	Ipotizzando che il 10% degli utenti possiede 2 carte registrate e che l'1% degli utenti abbiano registrato una carta di pagamento già registrata da un altro utente.
Metodo di Pagamento	R	11.000	Considerando che il 10% degli utenti possiede 2 carte registrate.
Nazionalità	R	10.000	Stesso volume di Utente.
Sottoscrizione	R	10.000	Stesso volume di Utente.
Offerta funzionalità	R	20.000	Ipotizzando 2 funzionalità aggiuntive inserite in media in ciascun abbonamento.
Abilitazione	R	10.000	Stesso volume di Abbonamento.
Restrizione Pacchetto	R	50	Ipotizzando che un pacchetto non è fruibile in media in 10 Stati.

10.1.3 Tavola degli Accessi

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Utente	E	Scrittura	Inserimento dati nuovo Utente.	1
2	Nazionalità	R	Scrittura	Inserimento relazione.	1
3	Carta di Pagamento	E	Scrittura	Inserimento dati Carta di Pagamento.	1
4	Metodo di Pagamento	R	Scrittura	Inserimento relazione.	1
5	Abbonamento	E	Scrittura	Creazione Abbonamento associato all'utente.	1
6	Sottoscrizione	R	Scrittura	Inserimento relazione.	1
7	Pacchetto	E	Lettura	Controllo se l' <i>età minima</i> del pacchetto selezionato è coerente con l'età del nuovo utente.	1

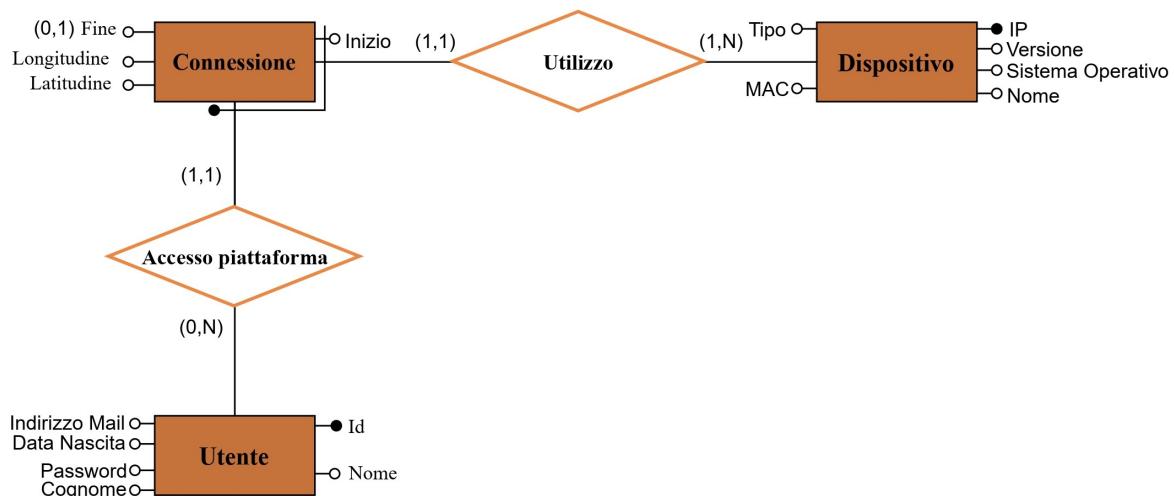
8	Restrizione Pacchetto	R	Lettura	Controllo sull'eventuale restrizione del pacchetto nello Stato di nazionalità del nuovo utente.	1
9	Abilitazione	R	Scrittura	Inserimento relazione.	1
10	Offerta funzionalità	R	Scrittura	Aggiunta in media di 2 funzionalità per ciascun abbonamento.	2
Totale Accessi per operazione = 11 (20 operazioni elementari)					
Totale Accessi giornalieri = 550 (1.000 operazioni elementari)					

10.2 Login alla Piattaforma

Il login da parte di un utente alla piattaforma è una delle operazioni che viene svolta con maggiore frequenza ed è dunque fondamentale per quanto riguarda la progettazione del servizio. Questa operazione consiste nell'andare a stabilire una connessione, identificata da un dispositivo, con il servizio, che permetterà successivamente di richiedere l'erogazione, e la conseguente visualizzazione da parte dell'utente, di un determinato contenuto multimediale. Per poter effettuare la procedura di login, un utente deve essersi precedentemente registrato alla piattaforma. Di conseguenza l'implementazione dovrà tenere conto della possibilità che un utente non abbia ancora svolto tale operazione.

Input	Output	Frequenza
Indirizzo Mail, Password, Attributi di Dispositivo , Latitudine e Longitudine	idUtenteLoggato	10.000 al giorno

10.2.1 Sezione Diagramma Interessato



10.2.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Clienti			
Utente	E	10.000	Ipotizzato.
Connessione	E	10.800.000	Ipotizzando che in media un utente si connette alla piattaforma una volta al giorno. Considerando le connessioni degli ultimi 3 anni si hanno: 10.000 utenti * 30 giorni = 300.000 connessioni al mese, 300.000 * 12 mesi = 3.600.000 all'anno, 3.600.000 * 3 anni = 10.800.000 connessioni totali.
Dispositivo	E	20.000	In media un utente si connette alla piattaforma utilizzando 2 diversi dispositivi.
Accesso piattaforma	R	10.800.000	Stesso volume di Connessione.
Utilizzo	R	10.800.000	Stesso volume di Connessione.

10.2.3 Tavola degli Accessi

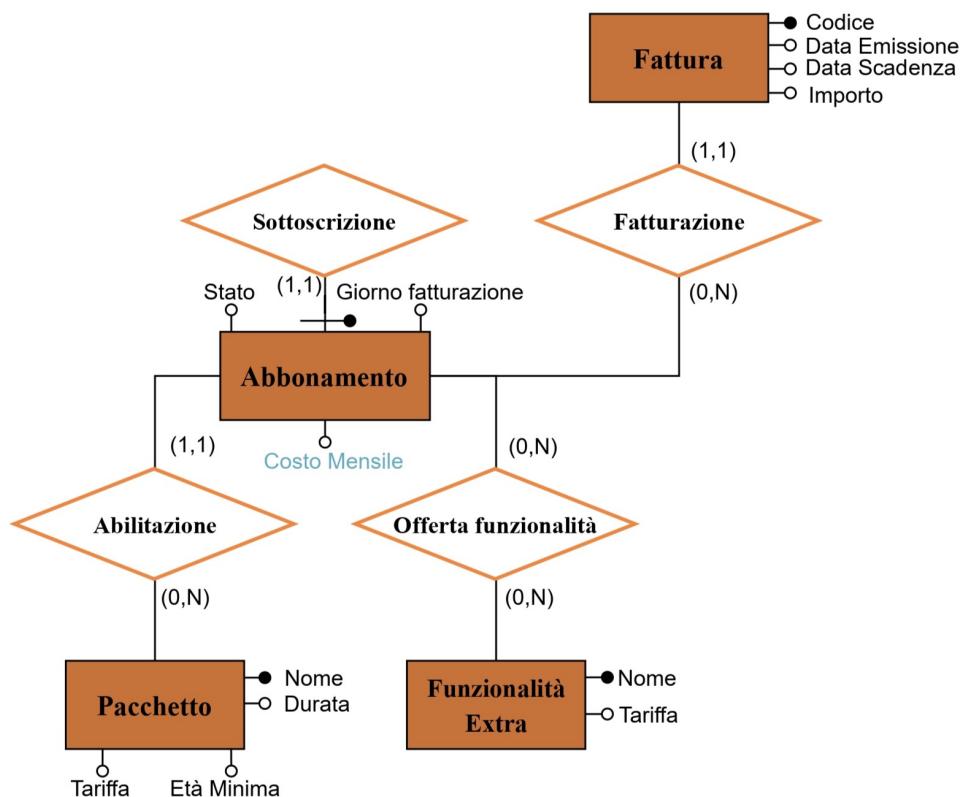
Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Utente	E	Lettura	Controllo sui dati Mail e Password inseriti.	1
2	Connessione	E	Lettura	Controllo possibile connessione in corso.	1
3	Dispositivo	E	Scrittura	Inserimento dati relativi al dispositivo di accesso.	1
4	Connessione	E	Scrittura	Inserimento entità.	1
5	Utilizzo	R	Scrittura	Inserimento relazione che associa il dispositivo alla connessione.	1
6	Accesso Piattaforma	R	Scrittura	Inserimento relazione che associa l'utente alla connessione.	1
Totale Accessi per operazione = 6 (10 operazioni elementari)					
Totale Accessi giornalieri = 60.000 (100.000 operazioni elementari)					

10.3 Emissione Fatture

L'operazione di emissione delle fatture consiste nell'andare a controllare, con cadenza giornaliera, gli abbonamenti che scadono i quel determinato giorno del mese ed emettere una fattura a carico dell'utente, per permettere a quest'ultimo di saldare il proprio conto con la piattaforma e continuare ad usufruire del servizio. Questa operazione viene ritenuta significativa perché è molto frequente e molto interessante per la valutazione delle prestazioni del sistema, nonché per quanto riguarda l'aspetto finanziario del business.

Input	Output	Frequenza
giornoFatturazione		1 al giorno

10.3.1 Sezione Diagramma Interessato



10.3.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Clienti			
Abbonamento	E	10.000	Stesso volume di Utente.
Pacchetto	E	5	La piattaforma supporta 5 modalità di abbonamento.
Funzionalità Extra	E	10	Ipotizzato.
Fattura	E	360.000	Ipotizzando solo le fatture degli ultimi 3 anni e considerando una fatturazione mensile si hanno 10.000 abbonamenti mensili attivi * 12 mesi * 3 anni = 360.000.
Fatturazione	R	360.000	Stesso volume di Fattura.
Offerta funzionalità	R	20.000	Ipotizzando 2 funzionalità aggiuntive inserite in media in ciascun abbonamento.
Abilitazione	R	10.000	Stesso volume di Abbonamento.

10.3.3 Tavola degli Accessi

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Abbonamento	E	Lettura	Controllo sulle fatture che hanno come <i>giorno fatturazione</i> il giorno corrente. Ipotizzando 10.000 abbonamenti attivi, in media al giorno saranno emesse 333 fatture.	333
2	Abilitazione	R	Lettura	Lettura relazione per ricavare il pacchetto abilitato.	333
3	Pacchetto	E	Lettura	Lettura entità per ricavare la <i>tariffa</i> di ciascun pacchetto, ipotizzando che su 333 abbonamenti ogni pacchetto sia stato abilitato almeno 1 volta.	5
4	Offerta Funzionalità	R	Lettura	Lettura delle funzionalità aggiunte in ciascun abbonamento, in media 2 per ciascuno.	666

5	Funzionalità Extra	E	Lettura	Ipotizzando che su 666 funzionalità extra aggiunte nei 333 abbonamenti da fatturare, sia stata inclusa almeno una volta ciascuna funzionalità disponibile nella piattaforma, servirà leggere tutte le istanze per ricavare la <i>tariffa</i> di ciascuna.	10
6	Fattura	R	Scrittura	Inserimento fattura con relativo importo calcolato con i valori letti.	333
7	Fatturazione	R	Scrittura	Inserimento relazione.	333
Totale Accessi al giorno = 2.013 (2.679 operazioni elementari)					

10.3.4 Valutazione della Ridondanza Costo Mensile in Abbonamento

Dato l'elevato numero di operazioni che vengono effettuate per ricavare la tariffa di ogni abbonamento per poter generare la fattura associata, è stata valutata l'introduzione di un attributo ridondante nell'entità *Abbonamento* chiamato *costo mensile*. Questo attributo viene inserito contestualmente alla registrazione dell'account e creazione dell'abbonamento e deve essere aggiornato ogni qualvolta viene aggiunta o eliminata una funzionalità extra presente nell'abbonamento.

Tavole degli Accessi con Ridondanza

Tabella 10.10: Tavola degli accessi con l'attributo ridondante *costo mensile*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Abbonamento	E	Lettura	Controllo sulle fatture che hanno come <i>giorno fatturazione</i> il giorno stesso dell'operazione. Ipotizzando 10.000 abbonamenti attivi, in media al giorno saranno emesse 333 fatture.	333
2	Fattura	E	Scrittura	Inserimento fattura con relativo importo <i>costo mensile</i> .	333
3	Fatturazione	R	Scrittura	Inserimento relazione.	333
Totale Accessi al giorno = 999 (1.665 operazioni elementari)					

L'operazione Emissione Fatture ha costo giornaliero di $n^T = 2.679$ operazioni elementari. A seguito dell'introduzione della ridondanza il numero di operazioni elementari per l'operazione è di $n_{RID}^T = 1.665$. Quindi il numero di operazioni risparmiate per eseguire l'operazione è di $\Delta = 1.014$ operazioni elementari.

L'operazione di aggiornamento della ridondanza avviene in due momenti:

- Durante la registrazione di un nuovo utente, che comporta l'aggiunta di 100 operazioni elementari giornaliere da svolgere in più.
- Durante l'aggiunta o l'eliminazione di una funzionalità extra, che comporta l'aggiunta di 100 operazioni elementari giornaliere da svolgere in più.

Tabella 10.11: Tavola degli accessi per la registrazione di un nuovo utente con l'attributo ridondante *costo mensile*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Utente	E	Scrittura	Inserimento dati nuovo Utente.	1
2	Nazionalità	R	Scrittura	Inserimento Nazionalità.	1
3	Carta di Pagamento	E	Scrittura	Inserimento dati Carta di Pagamento.	1
4	Metodo di Pagamento	R	Scrittura	Inserimento relazione.	1
5	Pacchetto	E	Lettura	Controllo se l' <i>età minima</i> del pacchetto selezionato è coerente con l'età del nuovo utente e lettura della tariffa.	1
6	Restrizione Pacchetto	R	Scrittura	Controllo sull'eventuale restrizione del pacchetto nello Stato di nazionalità del nuovo utente.	1
7	Funzionalità Extra	E	Lettura	Lettura della <i>tariffa</i> per le funzionalità extra aggiunte all'abbonamento.	2
8	Abbonamento	E	Scrittura	Creazione abbonamento associato all'utente compreso di <i>costo mensile</i> .	1
9	Sottoscrizione	R	Scrittura	Inserimento relazione.	1
10	Abilitazione	R	Scrittura	Inserimento dati nuovo utente.	1
11	Offerta Funzionalità	R	Scrittura	Aggiunta in media di 2 funzionalità per ciascun abbonamento.	2
Totale Accessi per operazione = $11 + 2$ (20 + 2 operazioni elementari)					
Totale Accessi giornalieri = $550 + 100$ (1.000 + 100 operazioni elementari)					

Nella registrazione di un nuovo utente con l'introduzione della ridondanza dovranno essere effettuate 2 operazioni elementari in più per ogni registrazione. Giornalmente saranno quindi, considerando una media di 50 operazioni, 100 operazioni elementari in più.

Tabella 10.12: Tavola degli accessi per l'aggiunta o eliminazione di una funzionalità extra in un abbonamento con l'attributo ridondante *costo mensile* con frequenza giornaliera di 25 operazioni.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Funzionalità Extra	E	Lettura	Lettura della <i>tariffa</i> della funzionalità da aggiungere o da eliminare.	1
2	Abbonamento	E	Lettura	Lettura di <i>costo mensile</i> precedente.	1
3	Abbonamento	E	Scrittura	Aggiornamento di <i>costo mensile</i> .	1
Totale Accessi per operazione = 3 (4 operazioni elementari)					
Totale Accessi giornalieri = 75 (100 operazioni elementari)					

Il numero totale di operazioni elementari giornaliere per mantenere aggiornato l'attributo ridondante ammonta a $n^A = 200$.

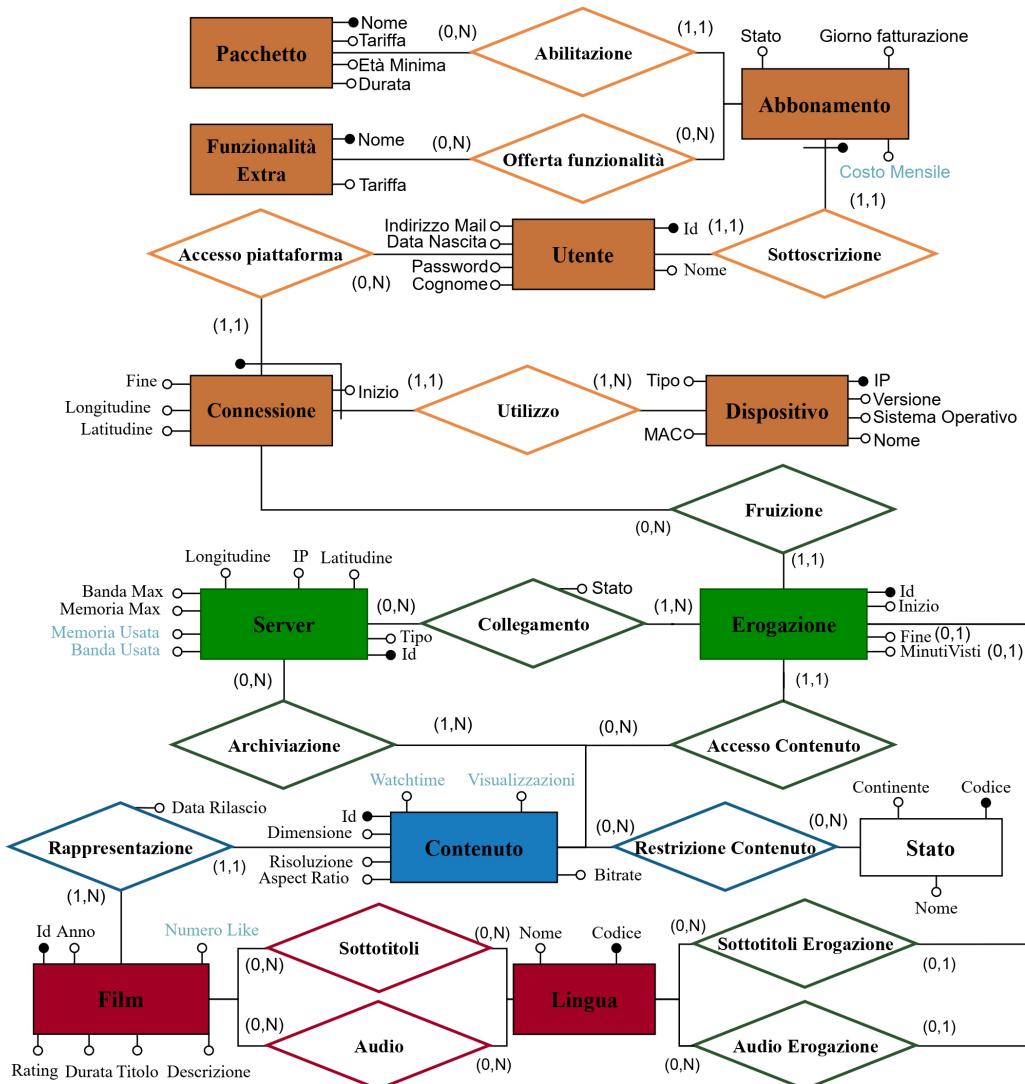
La ridondanza risulta conveniente poiché il numero di operazioni effettuate in presenza della ridondanza è di $n_{RID}^T + n^A = 1.865$ operazioni elementari che è inferiore alle $n^T = 2.679$ effettuate in assenza dell'attributo *costo mensile*.

10.4 Richiesta Visualizzazione Contenuto Video

La seguente operazione permette ad un utente, connesso al servizio, di richiedere di visualizzare un determinato contenuto video. L'operazione consiste nel creare un'erogazione del suddetto contenuto andando a selezionare il server in grado di fornire la migliore qualità di streaming possibile. L'operazione di selezione del server è stata ampiamente affrontata nella sezione 5.3.4. La frequenza di utilizzo di questa operazione è stata ipotizzata di una volta al giorno per ciascun utente, dunque di 10.000 volte al giorno.

Input	Output	Frequenza
Id (Contenuto), Id (Utente), Codice Audio(Lingua), Codice Sottotitoli(Lingua)		10.000 al giorno

10.4.1 Sezione Diagramma Interessato



10.4.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Contenuti			
Sottotitoli	R	6.000	Ipotizzando che un film in media abbia i sottotitoli disponibili in 15 lingue, $15 * 400 = 6.000$.
Audio	R	3.200	Ipotizzando che un film sia disponibile in media in 8 lingue, $8 * 400 = 3.200$.
Area Clienti			
Connessione	E	10.800.000	Ipotizzando che in media un utente si connette alla piattaforma una volta al giorno. Considerando le connessioni degli ultimi 3 anni si hanno: $10.000 \text{ utenti} * 30 \text{ giorni} = 300.000 \text{ connessioni al mese}$, $300.000 * 12 \text{ mesi} = 3.600.000 \text{ all'anno}$, $3.600.000 * 3 \text{ anni} = 10.800.000 \text{ connessioni totali}$.
Abbonamento	E	10.000	Stesso volume di Utente.
Accesso piattaforma	R	10.800.000	Stesso volume di Connessione.
Offerta funzionalità	R	20.000	Ipotizzando 2 funzionalità aggiuntive inserite in media in ciascun abbonamento.
Abilitazione	R	10.000	Stesso volume di Abbonamento.
Area Formati			
Contenuto	E	2.000	Ipotizzando che ogni film in media abbia 5 contenuti diversi che lo rappresentano, $5 * 400 = 2.000$ contenuti.
Restrizione Contenuto	R	20.000	Ipotizzando che un contenuto in media non sia fruibile in 10 Paesi, $10 * 2.000 = 20.000$.
Rappresentazione	R	2.000	Stesso volume di Contenuto.
Area Streaming			
Server	E	30	Ipotizzato.
Erogazione	E	10.800.000	Considerando la precedente ipotesi sulle connessioni giornaliere medie di 10.000, ipotizziamo che ad ogni connessione si richieda l'erogazione di un contenuto. Considerando le erogazioni degli ultimi 3 anni saranno lo stesso volume delle connessioni.
Archiviazione	R	7.500	Ipotizzando che ogni server contenga in media 250 contenuti, $250 * 30 = 7.500$.
Accesso Contenuto	R	10.800.000	Stesso volume di Erogazione.

Collegamento	R	11.016.000	Ipotizzando che per ogni erogazione ci sia un collegamento e che ogni 50 erogazioni ci sia un cambio di server. $10.800.000 + 10.800.000 / 50 = 11.016.000$.
Fruizione	R	10.800.000	Stesso volume di Erogazione.
Audio Erogazione	R	9.720.000	Ipotizzando che su tutte le erogazioni il 90% siano state visualizzate con un audio, mentre il restante 10% siano film muti.
Sottotitoli Erogazione	R	4.320.000	Ipotizzando che su tutte le erogazioni il 40% siano state visualizzate con dei sottotitoli.

10.4.3 Tavola degli Accessi

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Abbonamento	E	Lettura	Controllo <i>stato</i> abbonamento.	1
2	Abilitazione	R	Lettura	Lettura pacchetto abilitato e controllo possibilità di streaming del contenuto.	1
3	Offerta Funzionalità	R	Lettura	Lettura funzionalità extra aggiunte e controllo possibilità di streaming del contenuto.	2
4	Erogazione	E	Lettura	Lettura erogazioni dell'utente e controllo possibilità di streaming, ipotizzando in media 2 erogazioni attive.	2
5	Restrizione Contenuto	R	Lettura	Controllo possibilità di streaming.	1
6	Rappresenta- zione	R	Lettura	Lettura dell'identificatore del film relativo al contenuto.	1
7	Sottotitoli	R	Lettura	Controllo disponibilità sottotitoli.	1
8	Audio	R	Lettura	Controllo disponibilità traccia audio.	1
9	Accesso Piattaforma	R	Lettura	Lettura della relazione che permette di trovare la connessione attiva dell'utente.	1
10	Connessione	E	Lettura	Lettura degli attributi della connessione trovata.	1
11	Erogazione	E	Scrittura	Inserimento entità.	1

12	Audio Erogazione	R	Scrittura	Inserimento relazione.	1
13	Sottotitoli Erogazione	R	Scrittura	Inserimento relazione.	1
14	Fruizione	R	Scrittura	Inserimento relazione che associa la connessione dell'utente all'erogazione appena creata.	1
15	Archiviazione	R	Lettura	In media in archivio un server ha 250 contenuti su 2.000 disponibili. Avendo 30 server, in media un contenuto è archiviato in 4 server. Da queste 4 archiviazioni ricavo i 4 server che saranno da confrontare per la scelta di quello con costo streaming minore.	4
16	Server	E	Lettura	È necessario scegliere il server più opportuno, per farlo devono esser lette le posizioni geografiche dei server che possiedono il contenuto per calcolare il costo in termini di distanza dal punto di connessione dell'utente.	4
17	Collegamento	R	Lettura	Per ogni server trovato si dovrà stabilire qual'è il carico attuale in quel momento per poter aggiungere questo fattore al costo di streaming di quel server. In media ci sono 10.000 richieste di visualizzazioni contenuti al giorno, è possibile ipotizzare una distribuzione omogenea su ogni server, dunque ogni server avrà 10.000 collegamenti / 30 server = 333 collegamenti totali giornalieri per ciascun server. Di queste si ipotizza che, in media, contemporaneamente, possano esserci 100 collegamenti su un server. Dunque occorre controllare 100 collegamenti * 4 server= 400.	400
18	Accesso Contenuto	R	Lettura	Lettura relazione.	400

19	Contenuto	E	Lettura	Per ogni contenuto devo ricavare il <i>bitrate</i> per calcolare il costo streaming, si ipotizza che delle 400 erogazioni, i contenuti differenti siano 300.	300
20	Contenuto	E	Lettura	Lettura <i>bitrate</i> del contenuto da visualizzare.	1
21	Collegamento	R	Scrittura	Una volta stabilito il server più opportuno si inserisce la relazione collegamento tra la nuova erogazione e il server trovato.	1
22	Accesso Contenuto	R	Scrittura	Una volta certi che il collegamento sia stato stabilito si crea la relazione che associa l'erogazione al contenuto.	1
Totale Accessi per operazione = 1.127 (1.133 operazioni elementari)					
Totale Accessi giornalieri = 11.270.000 (11.330.000 operazioni elementari)					

10.4.4 Valutazione Ridondanze Memoria Usata e Banda Usata in Server

Dato il costo molto elevato per questa operazione è stata valutata l'introduzione di due attributi ridondanti nell'entità *Server* chiamati *memoria usata* e *banda usata*. Il primo memorizza la banda pubblica, in Mbit/s, utilizzata rispetto alla propria *banda max*, mentre il secondo tiene conto della memoria di archiviazione, in GB, occupata dai contenuti presenti sul server rispetto alla propria *memoria max*.

Tavole degli Accessi con Ridondanza

Tabella 10.16: Tavola degli accessi con l'introduzione di attributi ridondanti *memoria usata* e *banda usata*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Abbonamento	E	Lettura	Controllo stato abbonamento.	1
2	Abilitazione	R	Lettura	Lettura pacchetto abilitato e controllo possibilità di streaming del contenuto.	1
3	Offerta Funzionalità	R	Lettura	Lettura funzionalità extra aggiunte e controllo possibilità di streaming del contenuto.	2
4	Erogazione	E	Lettura	Lettura erogazioni dell'utente e controllo possibilità di streaming, ipotizzando in media 2 erogazioni attive.	2
5	Restrizione Contenuto	R	Lettura	Controllo possibilità di streaming.	1
6	Rappresenta- zione	R	Lettura	Lettura dell'identificatore del film relativo al contenuto.	1
7	Sottotitoli	R	Lettura	Controllo disponibilità sottotitoli.	1
8	Audio	R	Lettura	Controllo disponibilità traccia audio.	1
9	Accesso Piattaforma	R	Lettura	Lettura della relazione che permette di trovare la connessione attiva dell'utente.	1
10	Connessione	E	Lettura	Lettura degli attributi della connessione trovata.	1
11	Erogazione	E	Scrittura	Inserimento entità.	1
12	Audio Erogazione	R	Scrittura	Inserimento relazione.	1
13	Sottotitoli Erogazione	R	Scrittura	Inserimento relazione.	1
14	Fruizione	R	Scrittura	Inserimento relazione che associa la connessione dell'utente all'erogazione appena creata.	1
15	Contenuto	E	Lettura	Lettura <i>bitrate</i> del contenuto da visualizzare.	1

16	Archiviazione	R	Lettura	In media in archivio un server ha 250 contenuti su 2.000 disponibili. Avendo 30 server, in media un contenuto è archiviato in 4 server. Da queste 4 archiviazioni ricavo i 4 server che saranno da confrontare per la scelta di quello con costo streaming minore.	4
17	Server	E	Lettura	È necessario scegliere il server più opportuno, per farlo devono esser lette le informazioni dei server in possesso del contenuto, compresi gli attributi ridondanti <i>memoria usata</i> e <i>banda usata</i> , per calcolare il costo di streaming relativo.	4
18	Collegamento	R	Scrittura	Una volta stabilito il server più opportuno si inserisce la relazione collegamento tra la nuova erogazione e il server trovato.	1
19	Accesso Contenuto	R	Scrittura	Una volta certi che il collegamento sia stato stabilito si crea la relazione che associa l'erogazione al contenuto.	1
Totale Accessi per operazione = 27 (33 operazioni elementari)					
Totale Accessi giornalieri = 270.000 (330.000 operazioni elementari)					

L'operazione Richiesta Visualizzazione Contenuto Video ha costo giornaliero di $n^T = 11.330.000$ operazioni elementari. A seguito dell'introduzione della ridondanza il numero di operazioni elementari per l'operazione è di $n_{RID}^T = 330.000$. Quindi il numero di operazioni risparmiate per eseguire l'operazione è di $\Delta = 11.000.000$ operazioni elementari.

Le operazioni di aggiornamento degli attributi ridondanti avvengono nella seguente maniera:

- L'aggiornamento dell'attributo *banda usata* viene effettuato quando viene aggiunto un nuovo collegamento, quando un collegamento già presente passa allo stato *interrotto* e quando termina un collegamento. La frequenza totale dei primi due eventi è di 10.200 volte al giorno. La frequenza dell'ultimo è di 10.000 volte al giorno. Dunque la frequenza totale giornaliera per quanto riguarda l'aggiornamento dell'attributo *banda usata* è di 20.200 esecuzioni al giorno.

- L'aggiornamento dell'attributo *memoria usata* viene effettuato contestualmente all'aggiunta o all'eliminazione di un'istanza della relationship *Archiviazione* che in media avviene in totale 200 volte al giorno.

Tabella 10.17: Tavola degli accessi per l'aggiornamento di *banda usata* quando viene aggiunto un nuovo collegamento, quando un collegamento già presente passa allo stato *interrotto*, o quando termina un collegamento.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Collegamento	R	Lettura	Lettura della relazione Collegamento che è stata aggiunta, eliminata o modificata.	1
2	Accesso Contenuto	R	Lettura	Lettura relazione.	1
3	Contenuto	E	Lettura	Lettura <i>bitrate</i> contenuto.	1
4	Server	E	Lettura	Lettura del server e dell'attributo <i>banda usata</i> precedente.	1
5	Server	E	Scrittura	Aggiornamento dell'attributo <i>banda usata</i> .	1
Totale Accessi per operazione = 5 (6 operazioni elementari)					
Totale Accessi giornalieri = 101.000 (121.200 operazioni elementari)					

Tabella 10.18: Tavola degli accessi per l'aggiornamento dell'attributo *memoria usata* a seguito di un aggiunta o eliminazione di un contenuto dal server.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Archiviazione	R	Lettura	Lettura della relazione che è appena stata aggiunta o deve essere eliminata.	1
2	Contenuto	E	Lettura	Lettura <i>dimensione</i> del contenuto da aggiungere o eliminare.	1
3	Server	E	Lettura	Lettura del server e dell'attributo <i>memoria usata</i> precedente.	1
4	Server	E	Scrittura	Aggiornamento dell'attributo <i>memoria usata</i> .	1
Totale Accessi per operazione = 4 (5 operazioni elementari)					
Totale Accessi giornalieri = 800 (1.000 operazioni elementari)					

L'aggiornamento degli attributi ridondanti ha un costo giornaliero di $n^A = 121.200 + 1.000 = 122.200$ operazioni elementari. Il costo totale giornaliero per la richiesta di un contenuto multimediale in presenza delle ridondanze ammonta a $n_{RID}^T + n^A = 452.200$ operazioni elementari.

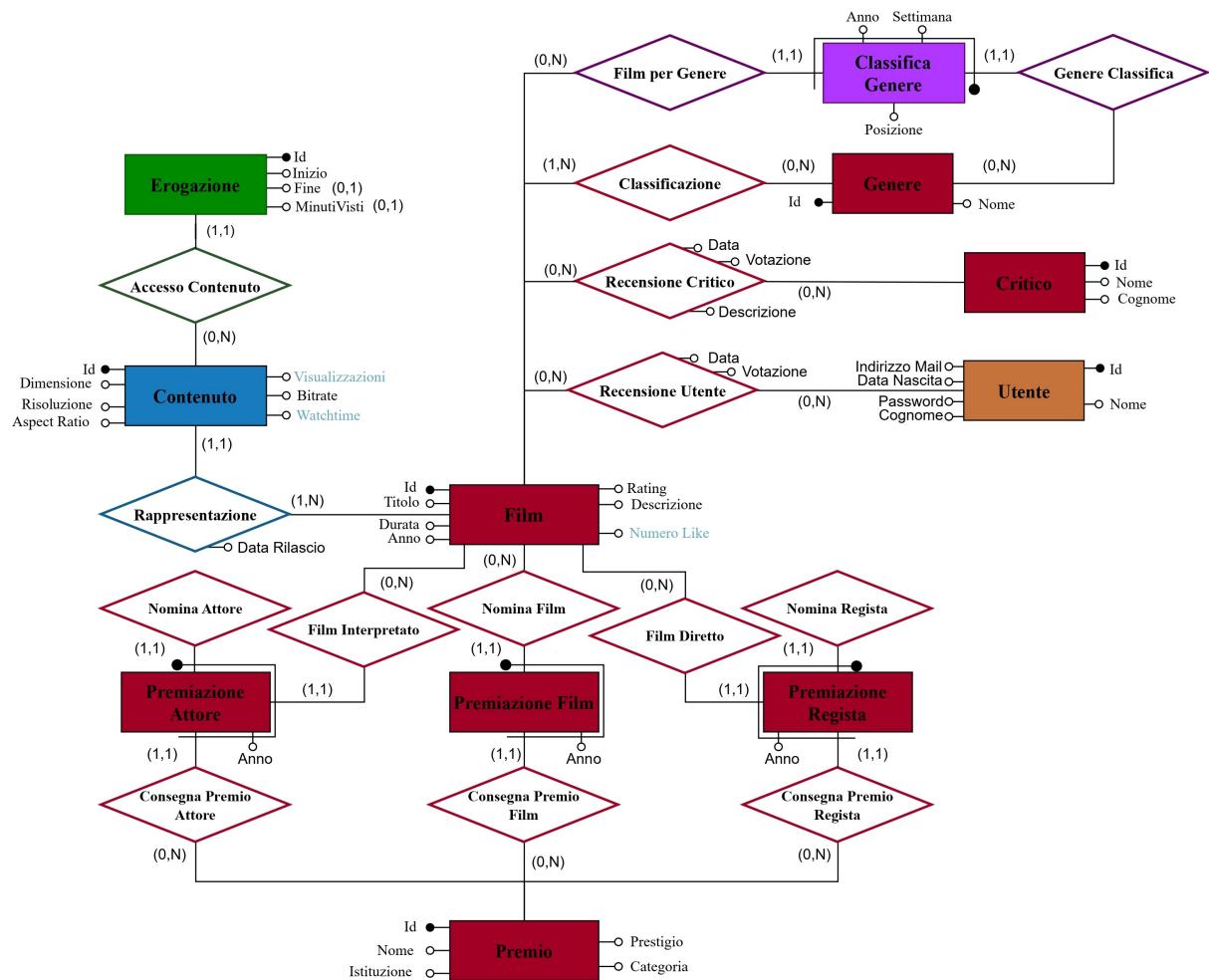
La ridondanza risulta conveniente poiché il numero di operazioni effettuate in presenza della ridondanza è di $n_{RID}^T + n^A = 452.200$ operazioni elementari giornaliere che è di gran lunga inferiore alle $n^T = 11.330.000$ effettuate in assenza degli attributi ridondanti.

10.5 Aggiornamento Rating e Classifica per Genere

Per l'implementazione della funzionalità di Raccomandazione Contenuti, FilmSphere necessita di stilare, con cadenza settimanale, delle classifiche in base al rating relativo ad ogni genere cinematografico presente sulla piattaforma. Prima di svolgere tale operazione, è necessario andare ad aggiornare il rating stesso di tutti i contenuti presenti, e solamente in un secondo momento sarà possibile andare a stilare le nuove classifiche. Di conseguenza l'operazione sarà composta da due sotto operazioni: l'aggiornamento del rating, e la creazione delle classifiche per genere. Per quanto riguarda l'immagazzinamento di queste ultime, il database mette a disposizione un'area analytics che permette di memorizzare tutte le classifiche per genere, in modo da avere uno storico settimanale sempre disponibile per essere consultato.

Input	Output	Frequenza
		Settimanale

10.5.1 Sezione Diagramma Interessato



10.5.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Contenuti			
Film	E	400	Ipotizzato.
Genere	E	20	Ipotizzato.
Premio	E	60	Ipotizzando che ogni anno vengano assegnati 20 premi per i film, 30 per gli attori e 10 per i registi.
Consegna Premio Film	R	30	Stesso volume di Premiazione Film.
Consegna Premio Attore	R	45	Stesso volume di Premiazione Attore.
Consegna Premio Regista	R	15	Stesso volume di Premiazione Regista.
Film Interpretato	R	45	Stesso volume di Premiazione Attore.
Film Diretto	R	15	Stesso volume di Premiazione Regista.
Nomina Film	R	30	Stesso volume di Premiazione Film.
Classificazione	R	800	Ipotizzando che un film in media appartiene a 2 generi diversi.
Recensione Critico	R	14.000	Ipotizzando che un film in media venga recensito dal 70% dei critici della piattaforma, dunque per ogni film ci saranno 35 recensioni di critici. $35 * 400 \text{ film} = 14.000$.
Recensione Utente	R	2.160.000	Ipotizzando che su 10.000 visualizzazioni giornaliere, circa il 20% di queste siano accompagnate anche da una recensione. $10.000 * 0.2 = 2.000$ recensioni al giorno, 60.000 al mese, 720.000 all'anno. Ipotizzando di conservare i dati degli ultimi 3 anni saranno presenti nel database circa $720.000 * 3 \text{ anni} = 2.160.000$ recensioni.
Area Formati			
Rappresentazione	R	2.000	Stesso volume di Contenuto.
Area Streaming			
Erogazione	E	10.800.000	Considerando la precedente ipotesi sulle connessioni giornaliere medie di 10.000, ipotizziamo che ad ogni connessione si richieda l'erogazione di un contenuto. Considerando le erogazioni degli ultimi 3 anni saranno lo stesso volume delle connessioni.

Accesso Contenuto	R	10.800.000	Stesso volume di Erogazione.
Area Analytics			
Classifica Genere	E	15.600	Avendo 20 generi, viene stilata la classifica dei 5 film con rating più alto per ciascun genere settimanalmente. $5 * 20$ generi = 100 record alla settimana. $100 * 52$ settimane * 3 anni = 15.600.
Film per Genere	R	15.600	Stesso volume di Classifica Genere.
Genere Classifica	R	15.600	Stesso volume di Classifica Genere.

10.5.3 Tavola degli Accessi

Tabella 10.21: Tavola degli accessi per aggiornamento *rating*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Erogazione	E	Lettura	Leggo i <i>minuti visti</i> del contenuto per tutte le erogazioni.	10.800.000
2	Accesso Contenuto	R	Lettura	Risalgo al contenuto visualizzato.	10.800.000
3	Rappresentazione	R	Lettura	Risalgo al film rappresentato da ciascun contenuto.	2.000
4	Film Interpretato	R	Lettura	Leggo le relazioni tra film e i premi attore.	45
5	Consegna Premio Attore	R	Lettura	Ricavo il premio.	45
6	Nomina Film	R	Lettura	Leggo le relazioni tra film e i premi film.	30
7	Consegna Premio film	R	Lettura	Ricavo il premio.	30
8	Film Diretto	R	Lettura	Leggo le relazioni tra film e i premi regista.	15
9	Consegna Premio Regista	R	Lettura	Ricavo il premio.	15

10	Premio	E	Lettura	Ricavo il <i>prestigio</i> dei premi ipotizzando che tutti i premi siano stati vinto almeno una volta da un film/regista/attore.	60
11	Recensione Critico	R	Lettura	Ricavo per ogni recensione dei critici la <i>votazione</i> .	14.000
12	Recensione Utente	R	Lettura	Ricavo per ogni recensione degli utenti la <i>votazione</i> .	2.160.000
13	Film	E	Scrittura	Per ciascun film ho tutti i dati necessari il calcolo del <i>rating</i> che viene aggiornato.	400
Totale Accessi per operazione = 23.776.640 (23.777.040 operazioni elementari)					

Tabella 10.22: Tavola degli accessi per creazione classifiche in base al genere.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Film	E	Lettura	Leggo tutti i <i>rating</i> aggiornati.	400
2	Classificazione	R	Lettura	Ricavo i generi per ciascun film, seguendo l'ipotesi che un film appartiene in media a 2 generi.	800
3	Classifica Genere	E	Scrittura	Inserimento dei 5 film con <i>rating</i> più alto per ogni genere, avendo 20 generi comporta l'inserimento di 100 record.	100
4	Film per Genere	R	Scrittura	Inserimento relazione.	100
5	Genere Classifica	R	Scrittura	Inserimento relazione.	100
Totale Accessi per operazione = 1.500 (1.800 operazioni elementari)					

10.5.4 Valutazione delle Ridondanze per Aggiornamento Rating

Il totale di Accessi per l'operazione di aggiornamento rating e la creazione delle classifiche per genere è di 23.778.140 accessi che valgono per 23.778.840 operazioni elementari con frequenza settimanale.

Dato l'elevato numero di operazioni necessarie per effettuare l'aggiornamento del rating di tutti i contenuti multimediali, è stata valutata l'introduzione di tre attributi ridondanti chiamati *visualizzazioni*, *watchtime* e *numero like*. I primi due attributi sono stati aggiunti all'entità *Contenuto* in modo da poter immagazzinare i dati dei diversi formati; il terzo attributo invece è stato aggiunto all'entità *Film*. Questi tre attributi permettono di andare a ridurre notevolmente la quantità di operazioni elementari necessarie per effettuare l'aggiornamento del rating.

L'attributo *visualizzazioni* immagazzina il numero totale di visualizzazioni per tale contenuto, l'attributo *watchtime* si occupa di memorizzare il numero totale di minuti per il quale un contenuto è stato visualizzato, mentre *numero like* indica il numero totale di recensioni positive da parte degli utenti che il film ha ricevuto.

Tavole degli Accessi con Ridondanza

Tabella 10.23: Tavola degli accessi per aggiornamento *rating* con attributi ridondanti.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Film	E	Lettura	Leggo tutti i film per ricavare il valore dell'attributo ridondante <i>numero like</i> .	400
2	Rappresentazione	R	Lettura	Leggo per ogni film tutte le occorrenze di Rappresentazione.	2.000
3	Contenuto	E	Lettura	Leggo tutti i contenuti per ricavare i valori degli attributi ridondanti <i>visualizzazioni</i> e <i>watchtime</i> .	2.000
4	Film Interpretato	R	Lettura	Leggo le relazioni tra film e i premi attore.	45
5	Consegna Premio Attore	R	Lettura	Ricavo il premio.	45
6	Nomina Film	R	Lettura	Leggo le relazioni tra film e i premi film.	30
7	Consegna Premio Film	R	Lettura	Ricavo il premio.	30
8	Film Diretto	R	Lettura	Leggo le relazioni tra film e i premi regista.	15
9	Consegna Premio Regista	R	Lettura	Ricavo il premio.	15

10	Premio	E	Lettura	Ricavo il prestigio dei premi ipotizzando che tutti i premi siano stati vinto almeno una volta da un film/regista/attore.	60
11	Recensione Critico	R	Lettura	Ricavo per ogni recensione dei critici la <i>votazione</i> .	14.000
12	Film	E	Scrittura	Scrittura del <i>rating</i> aggiornato per ogni film.	400
Totale Accessi per operazione = 19.040 (19.440 operazioni elementari)					

Dato che il costo della creazione delle classifiche settimanali per genere rimane invariato con o senza ridondanze, il confronto verrà fatto solo sull'aggiornamento del rating.

L'operazione Aggiornamento Rating avviene con cadenza settimanale ed ha un costo di $o^T = 23.777.040$ operazioni. Dunque il numero di operazioni giornaliere è $n^T = o^T/7 = 3.396.720$. Con l'introduzione degli attributi ridondanti le operazioni necessarie per l'aggiornamento del rating ammontano a $o_{RID}^T = 19.440$ operazioni elementari settimanali e, dunque, il numero di operazioni elementari giornaliere risulta $n_{RID}^T = o_{RID}^T/7 = 2.777$.

Le operazioni di aggiornamento degli attributi ridondanti avvengono nella seguente maniera:

- L'aggiornamento degli attributi *watchtime* e *visualizzazioni* viene effettuato contestualmente al termine di un'erogazione di un contenuto, che in media avviene 10.000 volte al giorno.
- L'aggiornamento dell'attributo *numero like* viene effettuato contestualmente all'aggiunta di una recensione positiva per un film, operazione effettuata in media 2.000 volte al giorno.

Tabella 10.24: Tavola degli accessi per l'aggiornamento di *watchtime* e *visualizzazioni*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Erogazione	E	Lettura	Leggo <i>minuti visti</i> del contenuto.	1
2	Accesso Contenuto	R	Lettura	Risalgo al contenuto visualizzato.	1
3	Contenuto	E	Lettura	Leggo il contenuto visualizzato e i suoi attributi.	1

4	Contenuto	E	Scrittura	Aggiorno i valori degli attributi ridondanti.	1
Totale Accessi per operazione = 4 (5 operazioni elementari)					
Totale Accessi giornalieri = 40.000 (50.000 operazioni elementari)					

Tabella 10.25: Tavola degli accessi per l'aggiornamento di *numero like*.

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Recensione Utente	R	Lettura	Leggo l'esito della recensione.	1
2	Film	E	Lettura	Leggo <i>numero like</i> precedenti per il film recensito.	1
3	Film	E	Scrittura	Aggiorno l'attributo ridondante <i>numero like</i> .	1
Totale Accessi per operazione = 2 (3 operazioni elementari)					
Totale Accessi giornalieri = 4.000 (6.000 operazioni elementari)					

L'aggiornamento degli attributi ridondanti ha un costo giornaliero di $n^A = 50.000 + 6.000 = 56.000$ operazioni elementari.

Il costo totale giornaliero per l'aggiornamento del rating ammonta a $n_{RID}^T + n^A = 2.777 + 56.000 = 58.777$ operazioni elementari.

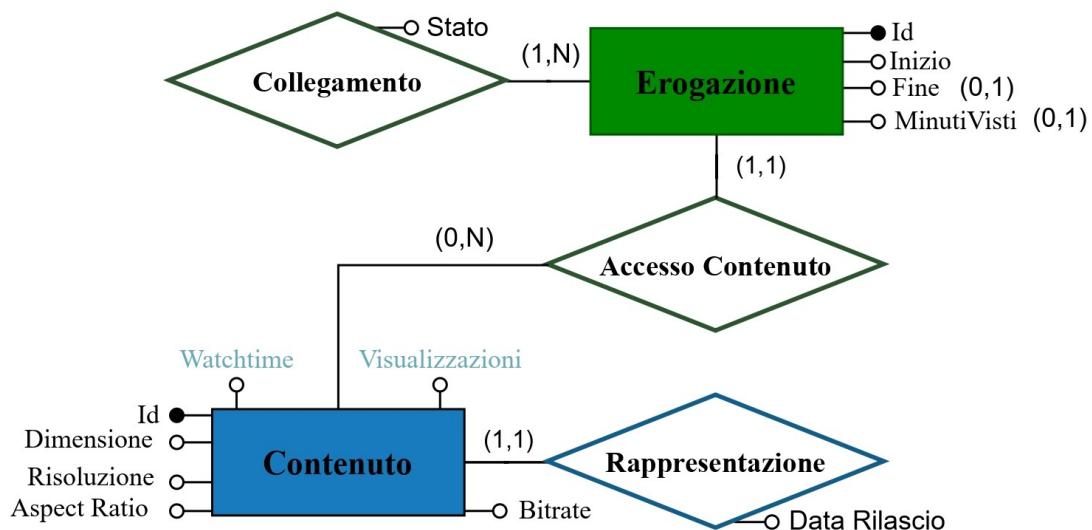
La ridondanza risulta conveniente poiché il numero di operazioni effettuate in presenza della ridondanza per l'aggiornamento del rating è di $n_{RID}^T + n^A = 58.777$ operazioni elementari giornaliere che è di gran lunga inferiore alle $n_T = 3.396.720$ effettuate in assenza degli attributi ridondanti.

10.6 Server più utilizzato per l'Erogazione di un Film

L'operazione permette di identificare quali sono i server che sono stati più utilizzati, nel corso di un periodo temporale specificato come input, per effettuare lo streaming di un determinato film, anch'esso fornito come input. Tale operazione risulta essere utile quando si va a valutare la aree in cui un determinato film viene richiesto maggiormente, per andare ad agire sui server della zona e incrementare l'archiviazione dei relativi contenuti.

Input	Output	Frequenza
Id (Film), Data Inizio Periodo, Data Fine Periodo	Id (Server)	Giornaliera

10.6.1 Sezione Diagramma Interessato



10.6.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Formati			
Rappresentazione	R	2.000	Stesso volume di Contenuto.
Area Streaming			
Accesso Contenuto	R	10.800.000	Stesso volume di Erogazione.
Collegamento	R	11.016.000	Ipotizzando che per ogni erogazione ci sia un collegamento e che ogni 50 erogazioni ci sia un cambio di server. $10.800.000 + 10.800.000 / 50 = 11.016.000$.

10.6.3 Tavola degli Accessi

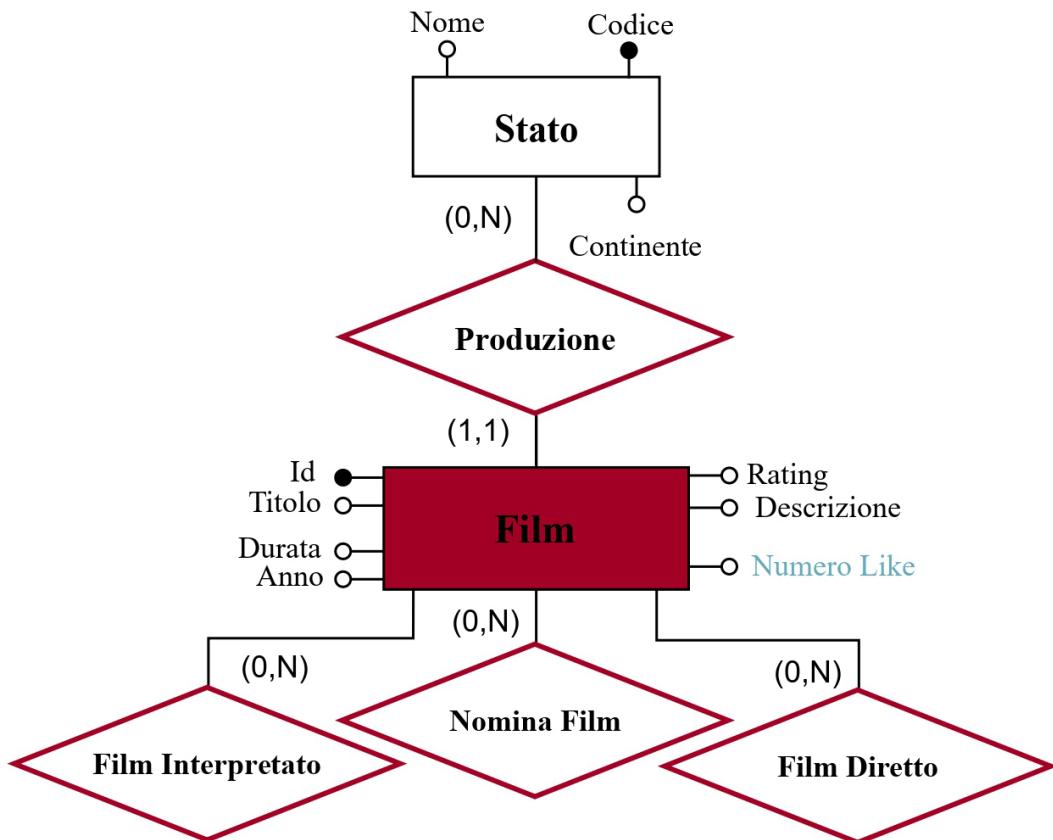
Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Rappresentazione	R	Lettura	Conoscendo l' <i>id</i> del film è possibile leggere direttamente le sue 5 rappresentazioni.	5
2	Accesso Contenuto	R	Lettura	Ipotizzando che il periodo scelto in media sia una settimana di tempo. Avendo 70.000 erogazioni di film alla settimana e che i film visti siano il 60% di quelli totali della piattaforma ovvero 240 film, ci saranno per un film in media 291 erogazioni settimanali.	291
3	Collegamento	R	Lettura	In media ogni 50 erogazioni si ha un cambiamento di server, perciò i collegamenti che dobbiamo leggere saranno $291 + 291/50 = 296$.	296
Totale Accessi per operazione = 592 (592 operazioni elementari)					

10.7 Stati con maggiori Film premiati

La seguente operazione permette di stilare una classifica degli Stati relativamente a quanti premi sono stati ottenuti in ogni categoria dai film prodotti nello Stato in questione. Dunque l'operazione restituirà, per ogni Stato di produzione che ha vinto almeno un premio, il numero di premi per la categoria registi, il numero di premi per la categoria film, il numero di premi per la categoria attori e il numero di premi totali, ognuno di essi affiancati da una classifica relativa realizzata attraverso l'impiego della window function *Dense Rank*.

Input	Output	Frequenza
	Premi per ogni Stato vincitore	Settimanale

10.7.1 Sezione Diagramma Interessato



10.7.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Contenuti			
Film Interpretato	R	45	Stesso volume di Premiazione Attore.
Film Diretto	R	15	Stesso volume di Premiazione Regista.
Nomina Film	R	30	Stesso volume di Premiazione Film.
Produzione	R	400	Stesso volume di Film.

10.7.3 Tavola degli Accessi

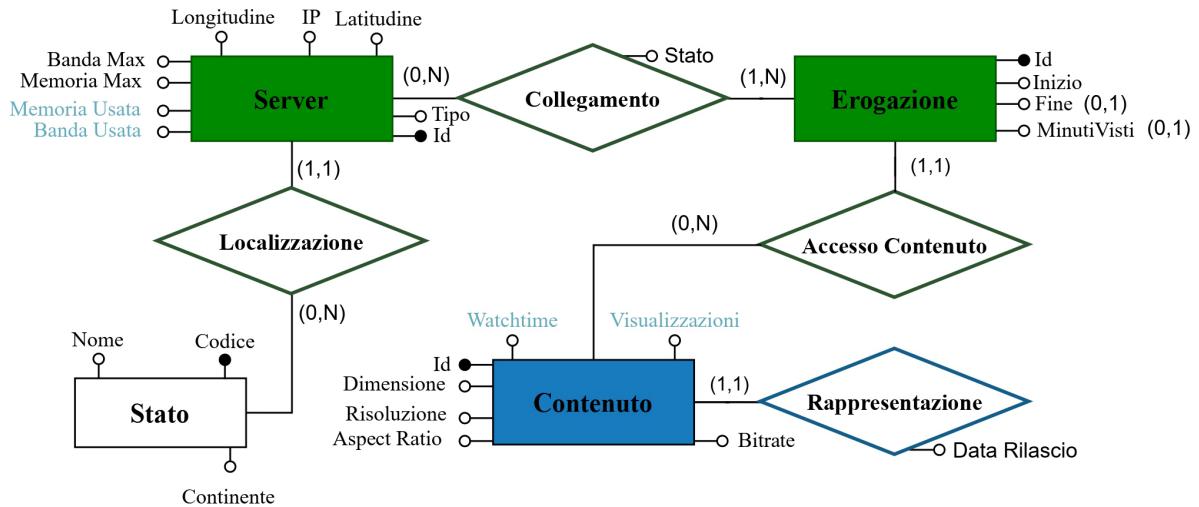
Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Film Interpretato	R	Lettura	Leggo tutte le premiazioni di attori associate ad un film.	45
2	Nomina Film	R	Lettura	Leggo tutte le premiazioni di film.	30
3	Film Diretto	R	Lettura	Leggo tutte le premiazioni di registi per film della piattaforma.	15
4	Produzione	R	Lettura	Mi ricavo lo stato di produzione di ciascuno dei film premiati, al massimo (se ogni film ha ricevuto non più di un premio in totale) questi saranno 90 e così sono in grado di stabilire in quale Stato sono stati prodotti tali film.	90
5	Stato	E	Lettura	Ricavo il <i>nome</i> dello Stato di produzione ipotizzando che il 50% dei film siano stati prodotti negli stessi stati	45
Totale Accessi per operazione = 225 (225 operazioni elementari)					

10.8 Distribuzione delle visualizzazioni di un Film

La seguente operazione permette di andare ad identificare la distribuzione della percentuale delle visualizzazioni di un determinato film fornito come input, nel corso di uno specificato lasso temporale, nei diversi continenti. Questa operazione può tornare molto utile per comprendere quali sono i continenti dove determinati film vengono visualizzati maggiormente e di conseguenza permettere di spostare i relativi contenuti nei server di tali continenti per favorirne la visualizzazione.

Input	Output	Frequenza
Id (Film), Data Inizio Periodo, Data Fine Periodo	Percentuale per ogni continente	1 volta al mese

10.8.1 Sezione Diagramma Interessato



10.8.2 Tavola dei Volumi Interessati

Concetto	Tipo	Volume	Considerazioni
Area Generale			
Stato	E	196	Il servizio sarà disponibile in 196 paesi.
Area Formati			
Rappresentazione	R	2.000	Stesso volume di Contenuto.
Area Streaming			
Server	E	30	Ipotizzato.
Erogazione	E	10.800.000	Considerando la precedente ipotesi sulle connessioni giornaliere medie di 10.000, ipotizziamo che ad ogni connessione si richieda l'erogazione di un contenuto. Considerando le erogazioni degli ultimi 3 anni saranno lo stesso volume delle connessioni.
Localizzazione	R	30	Stesso volume di Server.
Accesso Contenuto	R	10.800.000	Stesso volume di Erogazione.
Collegamento	R	11.016.000	Ipotizzando che per ogni erogazione ci sia un collegamento e che ogni 50 erogazioni ci sia un cambio di server. $10.800.000 + 10.800.000 / 50 = 11.016.000$.

10.8.3 Tavola degli Accessi

Id	Concetto	Tipo	Operazione	Considerazioni	Accessi
1	Rappresentazione	R	Lettura	Conoscendo già l' <i>id</i> del film, basta leggere le sue 5 rappresentazioni.	5
2	Accesso Contenuto	R	Lettura	Ipotizzando un lasso temporale della durata di un mese, in media ci sono sulla piattaforma 300.000 erogazioni di contenuti. Avendo 2.000 contenuti, di un contenuto sarà stata richiesta l'erogazione in media 150 volte. Dato che un film è rappresentato da 5 contenuti è necessario leggere 750 occorrenze di Accesso Contenuto.	750

3	Collegamento	R	Lettura	In media ogni 50 erogazioni abbiamo un cambiamento di server, perciò saranno $750/50 + 750 = 765$ collegamenti.	765
4	Localizzazione	R	Lettura	Ipotizzando che almeno un contenuto è stato richiesto su ciascun server della CDN, devo leggere tutte le localizzazioni di tutti i server.	30
5	Stato	E	Lettura	Ipotizzando i server siano stati posizionati per l'80% in stati diversi e il 20% di server nello stesso stato, diventano 24 Stati diversi.	24
Totale Accessi per operazione = 1.574 (1.574 operazioni elementari)					

Parte IV

Progettazione Logica

Descrizione Schema Logico

11.1 Schema Logico

Di seguito è riportata la descrizione dello schema logico, risultato del processo di traduzione del diagramma *entity-relationship* ristrutturato.

Per ogni relazione dello schema sono indicati gli attributi con il relativo tipo e le chiavi.

Per quanto riguarda i tipi degli attributi sono state seguite le seguenti linee guida:

- Per gli identificatori **unsigned int**: *uint*
- Per le percentuali **unsigned tiny int**: *utint*
- Per le durate **unsigned small int**: *usint*
- Per i timestamp **timestamp**: *ts*
- Per i valori boolean **boolean**: *bool*
- Per le dimensioni e i valori decimali **float**: *float*
- Per le stringhe **varchar(255)**: *vc(255)*
- Per gli identificatori di lunghezza fissa **char()**: *c()*

Stato(codice, *vc(255)*, *vc(255)*)
Artista (id, *uint*, *vc(255)*, *vc(255)*, *enum*, *date*, *date*, *date*)
Interpretazione (idArtista, idFilm, *uint*, *uint*, *vc(255)*)
Direzione (idArtista, idFilm, *uint*, *uint*)
Film (id, *uint*, *vc(255)*, *vc(1024)*, *usint*, *usint*, *utint*, *uint*, *c(2)*, *statoProduzione*)
PremiazioneAttore (idArtista, idFilm, *uint*, *uint*, *uint*, *usint*)
PremiazioneRegista (idArtista, idFilm, *uint*, *uint*, *uint*, *usint*)
PremiazioneFilm (idFilm, idPremio, *uint*, *uint*, *usint*)
Premio (id, *uint*, *vc(255)*, *vc(255)*, *utint*, *enum*)
Lingua (codice, *c(2)*, *vc(255)*)
Sottotitoli (idFilm, codiceLingua, *uint*, *c(2)*)

Audio (idFilm, codiceLingua)
 RecensioneUtente (idFilm, idUtente, data, votazione)
 Critico (id, nome, cognome)
 RecensioneCritico (idFilm, idCritico, data, votazione, descrizione)
 Genere (id, nome)
 Classificazione (idFilm, idGenere)
 Utente (id, nome, cognome, mail, password, nazionalità, dataDiNascita)
 Compatibilità (idFilm, idUtente, percentualeCompatibilità)
 Abbonamento (idUtente, stato, giornoFatturazione, costoMensile, nomePacchetto)
 Pacchetto (nome, durata, etaMinima, tariffa)
 FunzionalitaExtra (nome, tariffa)
 OffertaFunzionalita (idUtente, nomeFunzionalita)
 Fattura (codice, dataEmissione, dataScadenza, importo, idUtente, dataSaldo,
 numeroCartaSaldo)
 CartaDiPagamento (numero, scadenza, circuito, intestatario)
 MetodoDiPagamento (idUtente, numeroCarta)
 RestrizionePacchetto (codiceStato, nomePacchetto)
 Connessione (ipDispositivo, inizio, idUtente, fine, longitudine, latitudine)
 Dispositivo (ipAddress, versione, sistemaOperativo, nome, macAddress, tipo)
 Contenuto (id, dimensione, aspectRatio, bitrate, risoluzione, idFilm, watchtime,
 visualizzazioni, idFormato, idCodecAudio, idCodecVideo, dataRilascio)
 Formato (id, nome, versione, estensione)
 Codec (id, nome, compressione, tipologia, versione)
 RestrizioneContenuto (idContenuto, codiceStato)
 Server (id, memoriaUsata, bandaUsata, memoriaMax, bandaMax, tipo, ipAddress,
 latitudine, longitudine, localita)
 Archiviazione (idServer, idContenuto)
 Collegamento (idServer, idErogazione, stato)

Erogazione (id, uint, ts, ts, usint, uint, vc(15),
inizio, fine, minutiVisti, idContenuto, ipDispositivo,
ts, uint, c(2), c(2),
inizioConnessione, idUtente, codiceLinguaAudio, codiceLinguaSottotitoli)
ServerLog (id, uint, uint, ts, enum, enum, vc(1024))
ClassificaGenere (idFilm, uint, uint, usint, usint, usint,
settimana, anno, posizione)
ClassificaStato (idFilm, uint, c(2), usint, usint, usint, vc(255),
codiceStato, settimana, anno, posizione, risoluzione)
ClassificaPacchetto (idFilm, uint, vc(255), usint, usint, usint, vc(255),
nomePacchetto, settimana, anno, posizione, risoluzione)

11.2 Vincoli di Integrità

Per garantire l'integrità del database rispetto all'implementazione è necessario introdurre i seguenti vincoli di integrità.

11.2.1 Vincoli di Integrità Referenziale

- L'attributo *idArtista* di Interpretazione deve comparire come chiave primaria di Artista.
- L'attributo *idFilm* di Interpretazione deve comparire come chiave primaria di Film.
- L'attributo *statoProduzione* di Film deve comparire come chiave primaria di Stato.
- L'attributo *idArtista* di Direzione deve comparire come chiave primaria di Artista.
- L'attributo *idFilm* di Direzione deve comparire come chiave primaria di Film.
- L'attributo *idArtista* di PremiazioneAttore deve comparire come chiave primaria di Artista.
- L'attributo *idFilm* di PremiazioneAttore deve comparire come chiave primaria di Film.
- L'attributo *idPremio* di PremiazioneAttore deve comparire come chiave primaria di Premio.
- L'attributo *idArtista* di PremiazioneRegista deve comparire come chiave primaria di Artista.
- L'attributo *idFilm* di PremiazioneRegista deve comparire come chiave primaria di Film.
- L'attributo *idPremio* di PremiazioneRegista deve comparire come chiave primaria di Premio.
- L'attributo *idFilm* di PremiazioneFilm deve comparire come chiave primaria di Film.
- L'attributo *idPremio* di PremiazioneFilm deve comparire come chiave primaria di Premio.
- L'attributo *idFilm* di Sottotitoli deve comparire come chiave primaria di Film.
- L'attributo *codiceLingua* di Sottotitoli deve comparire come chiave primaria di Lingua.
- L'attributo *idFilm* di Audio deve comparire come chiave primaria di Film.
- L'attributo *codiceLingua* di Audio deve comparire come chiave primaria di Lingua.

- L'attributo *codiceLinguaAudio* di Erogazione deve comparire come chiave primaria di Lingua.
- L'attributo *codiceLinguaSottotitoli* di Erogazione deve comparire come chiave primaria di Lingua.
- L'attributo *idCritico* di RecensioneCritico deve comparire come chiave primaria di Critico.
- L'attributo *idFilm* di RecensioneCritico deve comparire come chiave primaria di Film.
- L'attributo *idGenere* di Classificazione deve comparire come chiave primaria di Genere.
- L'attributo *idFilm* di Classificazione deve comparire come chiave primaria di Film.
- L'attributo *nazionalita* di Utente deve comparire come chiave primaria di Stato.
- L'attributo *idUtente* di RecensioneUtente deve comparire come chiave primaria di Utente.
- L'attributo *idFilm* di RecensioneUtente deve comparire come chiave primaria di Film.
- L'attributo *nomePacchetto* di Abbonamento deve comparire come chiave primaria di Pacchetto.
- L'attributo *idUtente* di Abbonamento deve comparire come chiave primaria di Utente.
- L'attributo *idUtente* di OffertaFunzionalita deve comparire come chiave primaria di Abbonamento.
- L'attributo *nomeFunzionalita* di OffertaFunzionalita deve comparire come chiave primaria di FunzionalitaExtra.
- L'attributo *idUtente* di Compatibilita deve comparire come chiave primaria di Utente.
- L'attributo *idFilm* di Compatibilita deve comparire come chiave primaria di Film.
- L'attributo *idUtente* di Fattura deve comparire come chiave primaria di Utente.
- L'attributo *numeroCartaSaldo* di Fattura deve comparire come chiave primaria di CartaDiPagamento.
- L'attributo *numeroCarta* di MetodoDiPagamento deve comparire come chiave primaria di CartaDiPagamento.

- L'attributo *idUtente* di MetodoDiPagamento deve comparire come chiave primaria di Utente.
- L'attributo *codiceStato* di RestrizionePacchetto deve comparire come chiave primaria di Stato.
- L'attributo *nomePacchetto* di RestrizionePacchetto deve comparire come chiave primaria di Pacchetto.
- L'attributo *ipDispositivo* di Connessione deve comparire come chiave primaria di Dispositivo.
- L'attributo *idUtente* di Connessione deve comparire come chiave primaria di Utente.
- L'attributo *idFilm* di Contenuto deve comparire come chiave primaria di Film.
- L'attributo *idFormato* di Contenuto deve comparire come chiave primaria di Formato.
- L'attributo *idCodecAudio* di Contenuto deve comparire come chiave primaria di Codec.
- L'attributo *idCodecVideo* di Contenuto deve comparire come chiave primaria di Codec.
- L'attributo *idContenuto* di RestrizioneContenuto deve comparire come chiave primaria di Contenuto.
- L'attributo *codiceStato* di RestrizioneContenuto deve comparire come chiave primaria di Stato.
- L'attributo *localita* di Server deve comparire come chiave primaria di Stato.
- Gli attributi *ipDispositivo*, *inizioConnessione* e *idUtente* di Erogazione devono comparire come chiave primaria di Connessione.
- L'attributo *idContenuto* di Erogazione deve comparire come chiave primaria di Contenuto.
- L'attributo *idContenuto* di Archiviazione deve comparire come chiave primaria di Contenuto.
- L'attributo *idServer* di Archiviazione deve comparire come chiave primaria di Server.
- L'attributo *idErogazione* di Collegamento deve comparire come chiave primaria di Erogazione.
- L'attributo *idServer* di Collegamento deve comparire come chiave primaria di Server.
- L'attributo *idServer* di ServerLog deve comparire come chiave primaria di Server.

- L'attributo *idFilm* di ClassificaGenere deve comparire come chiave primaria di Film.
- L'attributo *nomeGenere* di ClassificaGenere deve comparire come chiave primaria di Genere.
- L'attributo *idFilm* di ClassificaStato deve comparire come chiave primaria di Film.
- L'attributo *codiceStato* di ClassificaStato deve comparire come chiave primaria di Stato.
- L'attributo *idFilm* di ClassificaPacchetto deve comparire come chiave primaria di Film.
- L'attributo *nomePacchetto* di ClassificaPacchetto deve comparire come chiave primaria di Pacchetto.

11.2.2 Vincoli di Integrità di Tupla

- L'attributo *dataDiNascita* di Utente deve essere di almeno 10 anni inferiore rispetto alla data di registrazione.
- L'attributo *prestigio* in Premio deve essere un numero compreso tra 1 e 5.
- L'attributo *percentualeCompatibilita* in Compatibilita deve essere un numero compreso tra 0 e 100.
- L'attributo *votazione* in RecensioneUtente deve essere un valore booleano che se positivo indica una recensione positiva, se negativo indica una recensione negativa.
- L'attributo *votazione* in RecensioneCritico deve essere un numero compreso tra 1 e 10.
- L'attributo *giornoFatturazione* in Abbonamento deve essere un'intero tra 1 e 28.
- L'attributo *etaMinima* in Pacchetto deve essere un numero compreso tra 14 e 18.
- L'attributo *dataScadenza* in Fattura deve essere maggiore dell'attributo *dataEmissione*.
- L'attributo *dataSaldo* in Fattura deve essere maggiore o uguale dell'attributo *dataEmissione*.
- Gli attributi *dataSaldo* e *numeroCartaSaldo* in Fattura devono essere entrambi nulli o entrambi non nulli.
- L'attributo *scadenza* in CartaDiPagamento deve essere maggiore della data attuale.
- L'attributo *latitudine* in Connessione deve essere un numero compreso tra -90 e +90.

- L'attributo *longitudine* in Connessione deve essere un numero compreso tra -180 e +180.
- L'attributo *inizio* in Connessione deve essere minore dell'attributo *fine*.
- L'attributo *ipApddress* in Dispositivo deve essere una stringa formata da 4 numeri interi compresi tra 0 e 255 separati da un punto, ad esempio: 89.201.17.32.
- L'attributo *macAddress* in Dispositivo deve essere una stringa formata da sei coppie di caratteri, dove ogni carattere è un valore esadecimale, separate dal simbolo "-", ad esempio: A3:B6:D4:32:45.
- L'attributo *latitudine* in Server deve essere un numero compreso tra -90 e +90.
- L'attributo *longitudine* in Server deve essere un numero compreso tra -180 e +180.
- L'attributo *ipAddress* in Server deve essere una stringa formata da 4 numeri interi compresi tra 0 e 255 separati da un punto, ad esempio: 89.201.17.32.
- L'attributo *bandaUsata* di Server deve essere minore o uguale dell'attributo *bandaMax*.
- L'attributo *memoriaUsata* di Server deve essere minore o uguale dell'attributo *memoriaMax*.
- L'attributo *fine* in Erogazione deve essere maggiore dell'attributo *inizio*.
- L'attributo *inizio* di Erogazione deve essere maggiore dell'attributo *inizioConnessione*.
- Gli attributi *fine* e *minutiVisti* in Erogazione devono essere entrambi nulli o entrambi non nulli.
- L'attributo *settimana* in ClassificaGenere deve essere un numero compreso tra 1 e 53.
- L'attributo *settimana* in ClassificaStato deve essere un numero compreso tra 1 e 53.
- L'attributo *settimana* in ClassificaPacchetto deve essere un numero compreso tra 1 e 53.
- L'attributo *genere* in Artista deve essere una stringa tra *uomo*, *donna* e *altro*.
- L'attributo *categoria* in Premio deve essere una stringa tra *attore*, *film* e *regista*.
- L'attributo *tipo* in Dispositivo deve essere una stringa tra *tablet*, *smartphone* e *desktop*.
- L'attributo *stato* di Collegamento deve essere una stringa tra *attivo*, *interrotto* e *terminato*.

- L'attributo *stato* in Abbonamento deve essere una stringa tra *attivo* e *interrotto*.
- L'attributo *tipologia* in Codec deve essere una stringa tra *audio* e *video*.
- L'attributo *risoluzione* in Contenuto deve essere una stringa tra *HD*, *FHD*, *UHD* e *FUHD*.
- L'attributo *criticità* in ServerLog deve essere una stringa tra *info*, *warning* e *error*.
- L'attributo *codice* in ServerLog deve essere una stringa tra *excessive load*, *excessive memory usage*, *server shutdown*, *server startup*, *server status* e *streaming error*.
- L'attributo *circuito* in CartaDiPagamento deve essere una stringa tra *visa*, *mastercard*, *maestro*, *american express*, *server status*, *discovery* e *jbc*.
- L'attributo *tipo* in Server deve essere una stringa tra *origin* e *edge*.

11.2.3 Vincoli di Integrità Generici

- L'attributo *idCodecAudio* di Contenuto deve comparire come chiave primaria di un record di Codec che presenti come attributo *tipologia* il valore *audio*.
- L'attributo *idCodecVideo* di Contenuto deve comparire come chiave primaria di un record di Codec che presenti come attributo *tipologia* il valore *video*.
- L'attributo *idPremio* di PremiazioneAttore deve comparire come chiave primaria di un record di Premio che presenti come attributo *categoria* il valore *attore*.
- L'attributo *idPremio* di PremiazioneRegista deve comparire come chiave primaria di un record di Premio che presenti come attributo *categoria* il valore *regista*.
- L'attributo *idPremio* di PremiazioneFilm deve comparire come chiave primaria di un record di Premio che presenti come attributo *categoria* il valore *film*.
- L'attributo *dataRilascio* in Contenuto deve essere minore della data attuale e maggiore o uguale dell'anno di rilascio del relativo Film.
- Per poter inserire un'istanza di PremiazioneRegista, occorre che il regista abbia diretto il film per cui è stato premiato.
- Per poter inserire un'istanza di PremiazioneAttore, occorre che l'attore abbia interpretato il film per cui è stato premiato.
- Per poter inserire un'istanza di PremiazioneRegista, PremiazioneAttore e PremiazioneFilm, occorre che il film sia stato pubblicato prima dell'attribuzione del premio.
- Per poter inserire un'istanza di RecensioneCritico, occorre che il film recensito sia stato pubblicato prima della pubblicazione della recensione.

- Per poter inserire un’istanza di RecensioneUtente, occorre che il suddetto utente abbia visualizzato il relativo contenuto almeno per il 25% della sua interezza.
- Una singola istanza di Erogazione può avvenire solamente attraverso un server alla volta, non è possibile avere più istanze di Collegamento con attributo *stato* attivo contemporaneamente per la stessa erogazione.
- Non deve essere consentito creare un’istanza di Erogazione relativa ad un Contenuto e ad un Utente per il quale tale Contenuto è ristretto nello Stato di registrazione.
- Non deve essere consentito creare un’istanza di Erogazione relativa ad una Connessione terminata.
- L’attributo *minutiVisti* di Erogazione deve essere minore o uguale della durata del contenuto visualizzato.
- Non deve essere consentito creare un’istanza di Collegamento tra un’Erogazione e un Server che non abbia il Contenuto richiesto in memoria, o che non abbia abbastanza banda disponibile per erogarlo.
- Non deve essere consentito eliminare un’istanza di Archiviazione, mentre tale contenuto si trova in un’Erogazione attiva.
- Non deve essere consentito creare un’istanza di Archiviazione tra un Contenuto e un Server che non abbia abbastanza memoria per contenerlo.
- Non deve essere consentito creare un’istanza di Erogazione quando il numero di erogazioni attive è uguale al numero massimo di erogazioni consentite dal piano di abbonamento o delle funzionalità extra dell’utente.
- La CartaDiPagamento utilizzata per saldare una Fattura deve essere un’istanza di MetodoDiPagamento per l’Utente che sta effettuando il saldo e non deve essere scaduta.
- Non deve essere consentito creare un’istanza di Abbonamento relativa ad un Pacchetto e ad un Utente per il quale tale Pacchetto è ristretto nello Stato di registrazione dell’utente e l’età dell’utente stesso non è conforme con l’attributo *etaMinima* del relativo Pacchetto.
- Non deve essere consentito creare un’istanza di OffertaFunzionalita tra un Utente e una FunzionalitaExtra che è già presente nel Pacchetto abilitato.
- Non deve essere consentito creare un’istanza di Erogazione che presenti attributi *codiceLinguaAudio* e *codiceLinguaSottotitoli* non presenti rispettivamente come istanze di Audio e Sottotitoli per il film relativo al contenuto erogato.

11.3 Analisi Dipendenze Funzionali e Normalizzazione

Di seguito è riportata l'analisi delle dipendenze funzionali per quanto riguarda ogni relazione della base di dati.

- Stato (codice, nome, continente)

codice → nome, continente
- Artista (id, nome, cognome, genere, dataDiNascita, dataDiMorte)

id → nome, cognome, genere, dataDiNascita, dataDiMorte
- Interpretazione (idArtista, idFilm, personaggio)

idArtista, idFilm → personaggio
- Direzione (idArtista, idFilm)

Non vi sono dipendenze funzionali non banali
- Film (id, titolo, descrizione, durata, annoProduzione, rating, numeroLike, statoProduzione)

id → titolo, descrizione, durata, annoProduzione, rating, numeroLike, statoProduzione
- PremiazioneAttore (idArtista, idFilm, idPremio, anno)

Non vi sono dipendenze funzionali non banali
- PremiazioneRegista (idArtista, idFilm, idPremio, anno)

Non vi sono dipendenze funzionali non banali
- PremiazioneFilm (idFilm, idPremio, anno)

Non vi sono dipendenze funzionali non banali
- Premio (id, nome, istituzione, prestigio, categoria)

id → nome, istituzione, prestigio, categoria
- Lingua (codice, nome)

codice → nome
- Sottotitoli (idFilm, codiceLingua)

Non vi sono dipendenze funzionali non banali
- Audio (idFilm, codiceLingua)

Non vi sono dipendenze funzionali non banali
- RecensioneUtente (idFilm, idUtente, data, votazione)

idFilm, idUtente → data, votazione

- Critico (id, nome, cognome)

id → nome, cognome
- RecensioneCritico (idFilm, idCritico, data, votazione, descrizione)

idFilm, idCritico → data, votazione, descrizione
- Genere (id, nome)

id → nome
- Classificazione (idFilm, idGenere)

Non vi sono dipendenze funzionali non banali
- Utente (id, nome, cognome, mail, password, nazionalità, dataDiNascita)

id → nome, cognome, mail, password, nazionalità, dataDiNascita

mail → id, nome, cognome, password, nazionalità, dataDiNascita
- Compatibilità (idFilm, idUtente, percentualeCompatibilita)

idFilm, idUtente → percentualeCompatibilita
- Abbonamento (idUtente, stato, giornoFatturazione, costoMensile, nomePacchetto)

idUtente → stato, giornoFatturazione, costoMensile, nomePacchetto
- Pacchetto (nome, durata, etaMinima, tariffa)

nome → durata, etaMinima, tariffa
- FunzionalitaExtra (nome, tariffa)

nome → tariffa
- OffertaFunzionalita (idUtente, nomeFunzionalita)

Non vi sono dipendenze funzionali non banali
- Fattura (codice, dataEmissione, dataScadenza, importo, idUtente, dataSaldo, numeroCartaSaldo)

codice → dataEmissione, dataScadenza, importo, idUtente, dataSaldo, numeroCartaSaldo
- CartaDiPagamento (numero, scadenza, circuito, intestatario)

numero → scadenza, circuito, intestatario
- MetodoDiPagamento (idUtente, numeroCarta)

Non vi sono dipendenze funzionali non banali
- RestrizionePacchetto (codiceStato, nomePacchetto)

Non vi sono dipendenze funzionali non banali

- Connessione (ipDispositivo, inizio, idUtente, fine, longitudine, latitudine)

ipDispositivo, inizio, idUtente → fine, longitudine, latitudine
- Dispositivo (ipAddress, versione, sistemaOperativo, nome, macAddress, tipo)

ipAddress → versione, sistemaOperativo, nome, macAddress, tipo
- Contenuto (id, dimensione, aspectRatio, bitrate, risoluzione, idFilm, watchtime, visualizzazioni, idFormato, idCodecAudio, idCodecVideo, dataRilascio)

id → dimensione, aspectRatio, bitrate, risoluzione, idFilm, watchtime, visualizzazioni, idFormato, idCodecAudio, idCodecVideo, dataRilascio
- Formato (id, nome, versione, estensione)

id → nome, versione, estensione
- Codec (id, nome, compressione, tipologia, versione)

id → nome, compressione, tipologia, versione
- RestrizioneContenuto (idContenuto, codiceStato)

Non vi sono dipendenze funzionali non banali
- Server (id, memoriaUsata, bandaUsata, memoriaMax, bandaMax, tipo, ipAddress, latitudine, longitudine, localita)

id → memoriaUsata, bandaUsata, memoriaMax, bandaMax, tipo, ipAddress, latitudine, longitudine, localita
- Archiviazione (idServer, idContenuto)

Non vi sono dipendenze funzionali non banali
- Collegamento (idServer, idErogazione, stato)

idServer, idErogazione → stato
- Erogazione (id, inizio, fine, minutiVisti, idContenuto, ipDispositivo, inizioConnessione, idUtente, codiceLinguaAudio, codiceLinguaSottotitoli)

id → inizio, fine, minutiVisti, idContenuto, ipDispositivo, inizioConnessione, idUtente, codiceLinguaAudio, codiceLinguaSottotitoli
- ServerLog (id, idServer, timestamp, criticita, codice, messaggio)

id → idServer, timestamp, criticita, codice, messaggio
- ClassificaGenere (idFilm, idGenere, settimana, anno, posizione)

idFilm, idGenere, settimana, anno → posizione
- ClassificaStato (idFilm, codiceStato, settimana, anno, posizione, risoluzione)

idFilm, codiceStato, settimana, anno → posizione, risoluzione

- ClassificaPacchetto (idFilm, nomePacchetto, settimana, anno, posizione, risoluzione)
 $\text{idFilm}, \text{nomePacchetto}, \text{settimana}, \text{anno} \rightarrow \text{posizione}, \text{risoluzione}$

Ne risulta che ogni relazione R presente si trova già in forma normale di Boyce-Codd dato che per ogni dipendenza funzionale non banale $X \rightarrow Y \in F^+$ si ha che X è superchiave della relazione R .

Funzioni Analytics

FilmSphere mette a disposizione diverse funzionalità lato server che permettono di analizzare i dati alla ricerca di informazioni utili a migliorare sia l'esperienza utente che l'efficienza della CDN.

12.1 Classifiche

Per valutare la popolarità dei contenuti presenti sulla piattaforma, FilmSphere mette a disposizione una funzionalità che permette di stilare delle classifiche settimanali relative a diversi fattori. Le suddette classifiche vengono inoltre immagazzinate all'interno del database per permetterne la fruizione in un secondo momento. Sono disponibili tre tipologie di classifiche: Classifica per Stato, Classifica per Pacchetto e Classifica per Genere.

12.1.1 Classifica per Stato

La classifica per Stato andrà a valutare i migliori film per ogni Stato in base alle loro visualizzazioni relative agli utenti che provengono da tale Stato. La posizione in classifica tiene conto sia delle visualizzazioni totali dall'aggiunta del contenuto alla data corrente, ma anche della percentuale di watchtime relativo a tali visualizzazioni secondo la formula:

$$Score = \frac{\text{Totale minuti visualizzati}}{\text{Durata Contenuto}}$$

Questa metrica permette non solo di comprendere quanto un determinato film è stato visualizzato, ma anche quanto è stato l'interesse riscontrato da parte degli spettatori nei confronti del film stesso.

Dunque più un contenuto è stato visualizzato nella sua interezza dagli utenti e più la sua posizione nella classifica è elevata.

Ogni film presente in classifica presenterà anche la risoluzione prediletta dagli utenti della piattaforma, per permettere di comprendere quale risoluzione viene maggiormente usata per la visualizzazione del film in questione.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni:

1. Crea un Result Set *FilmErogati* che contiene tutte le informazioni di ciascuna erogazione terminata presente nella piattaforma, estesa con dei dati aggiuntivi tra cui la nazionalità dell'utente che l'ha effettuata, la risoluzione del contenuto e la durata del film.

2. Crea un Result Set *WatchtimePerRisoluzione* che, a partire da *FilmErogati*, raggruppa il Result Set rispetto al film, Stato e risoluzione e calcola per ogni raggruppamento il totale di minuti visti.
3. Crea un Result Set *WatchtimeMaxPerFilmEStato* che a partire da *WatchtimePerRisoluzione* seleziona il watchtime massimo di visualizzazioni tra le diverse risoluzioni relative ad ogni film e Stato.
4. Crea un Result Set *RisoluzioniPerFilmEStato* che a partire da *WatchtimeMaxPerFilmEStato* associa a ciascun film e Stato la risoluzione in cui è stato visto più volte dagli utenti di quella nazionalità, e il numero della riga di ogni film e Stato utilizzando la window function *Row Number*, in modo da poter successivamente, in casi di ex aequo, selezionare solo una riga per raggruppamento.
5. Crea un Result Set *RisoluzioneUnicaPerFilmEStato* che a partire da *RisoluzioniPerFilmEStato* seleziona un'unica risoluzione se ve ne sono più di una che hanno lo stesso watchtime massimo per quel film e quello Stato.
6. Crea un Result Set *FilmStatoScore* che a partire dalla prima CTE *FilmErogati* calcola lo score di ciascun film considerando solo le visualizzazioni di utenti della stessa nazionalità.
7. Il Result Set finale effettua, tramite la window function *Dense Rank* sullo score, la classifica per ogni Stato e associa a ciascun film, la risoluzione in cui è stato visualizzato maggiormente tramite un join con il Result Set di *RisoluzioneUnicaPerFilmEStato*.

La procedura *CreazioneClassificaStato* permette di utilizzare la query sopra descritta per popolare la relazione *ClassificaStato* con la classifica relativa alla settimana corrente. Successivamente è possibile andare ad usufruire della classifica precedentemente generata attraverso la procedura *MostraClassificaStato*.

Per questa funzionalità è stato creato un Event che viene chiamato alla fine di ogni settimana in modo che generi la classifica in modo autonomo compresa dei dati della settimana appena conclusa.

12.1.2 Classifica per Pacchetto

La classifica per Pacchetto andrà a valutare le visualizzazioni dei film da parte degli utenti a seconda del pacchetto che hanno abilitato nel loro abbonamento. Per ogni pacchetto disponibile sarà stilata una classifica contenente tutti i film ordinati in base alla metrica definita nella seguente modalità.

$$Score = \frac{\text{Totale minuti visualizzati}}{\text{Durata Contenuto}}$$

Ogni film presente in classifica presenterà anche la risoluzione prediletta dagli utenti della piattaforma, per permettere di comprendere se effettivamente gli utenti che presentano un abbonamento che permette una visualizzazione maggiore effettivamente utilizzano o meno la possibilità di visualizzare contenuti con un'elevata risoluzione.

L'implementazione della funzionalità è stata gestita tramite una query analoga a quella utilizzata per generare la classifica per Stato. Al posto di raggruppare per lo Stato di nazionalità dell'utente che ha visualizzato un contenuto, si raggruppa per il pacchetto abilitato nell'abbonamento dell'utente.

La procedura *CreazioneClassificaPacchetto* permette di utilizzare la query sopra descritta per popolare la relazione *ClassificaPacchetto* con la classifica relativa alla settimana corrente.

Successivamente è possibile andare ad usufruire della classifica precedentemente generata attraverso la procedura *MostraClassificaPacchetto*.

Anche per questa funzionalità è stato creato un Event che viene chiamato alla fine di ogni settimana in modo che generi la classifica in modo autonomo compresa dei dati della settimana appena conclusa.

12.1.3 Classifica per Genere

La classifica per Genere va a valutare i 5 film con rating più alto per ogni genere presente sulla piattaforma. A differenza delle classifiche descritte precedentemente, che vengono valutate in base alla metrica definita negli appositi paragrafi, che tiene conto sia delle visualizzazioni totali, nonché della percentuale di visualizzazione dei contenuti stessi, le classifiche per Genere vengono calcolate in base al rating, un punteggio che viene associato ad ogni film e dipende da diversi fattori descritti largamente nella sezione dedicata (2.3.3). L'aggiornamento è settimanale e tiene conto dell'anno e del numero della settimana in cui è stata creata. Le suddette classifiche per genere non hanno solamente un impiego dal punto di vista analitico ma vengono utilizzate dal sistema di raccomandazione contenuti per andare a generare raccomandazioni personalizzate, come è stato ampiamente descritto nel relativo paragrafo (4.3.4).

La classifica viene generata tramite la window function *Dense Rank* in base al rating, calcolato precedentemente attraverso l'apposita operazione, di ciascun film e permette casi di pari merito inserendo più film nella stessa posizione per quel genere.

La procedura *CreazioneClassificaGenere* permette di utilizzare la query sopra descritta per popolare la relazione *ClassificaGenere* con la classifica relativa alla settimana corrente.

Successivamente è possibile andare ad usufruire della classifica precedentemente generata attraverso la procedura *MostraClassificaGenere*.

Per questa funzionalità è stato creato un Event che viene chiamato alla fine di ogni settimana in modo che vada ad aggiornare il rating dei film presenti sulla piattaforma e generi la classifica in modo autonomo con i rating aggiornati alla settimana appena conclusa.

12.2 Bilanciamento del Carico

Il sistema fornisce una funzionalità di bilanciamento del carico della CDN, implementata attraverso l'analisi dei log provenienti dai server della rete stessa.

I Server della CDN possono emettere i seguenti log, divisi in tre livelli di criticità:

- **Info:**

- *Server Startup*: log relativo all'avvio di un server.
- *Server Shutdown*: log relativo allo spegnimento di un server
- *Server Status*: log relativo allo stato della memoria e del carico del server, emesso ogni ora.

- **Warning:**

- *Excessive Load*: log emesso quando il carico di un server supera l' 80%.
- *Excessive Memory Usage*: log emesso quando la memoria occupata dai contenuti sul server supera il 90%.

- **Error:**

- *Streaming Error*: log emesso quando un server raggiunge il 95% di carico.

Tali log permetteranno di andare a valutare lo stato dell'intera rete e permetteranno di andare ad agire in caso di sovraccarico.

FilmSphere mette a disposizione una serie di funzionalità con la finalità di proporre piani di archiviazione personalizzati che permettono di andare a ridurre i tempi di latenza e ridurre i potenziali sovraccarichi dei server della CDN.

12.2.1 Gestione Excessive Load

GestioneExcessiveLoad è una procedura che richiede come input l'id del server e mostra per ogni erogazione in corso su tale server, il server più opportuno in quel momento su cui è possibile trasferire il collegamento.

La valutazione del server su cui spostare l'erogazione viene valutata attraverso la considerazione del costo di streaming, discusso ampiamente nel paragrafo relativo alla scelta del server per effettuare un'erogazione (5.3.4).

Questa funzionalità diventa fondamentale nel caso in cui si verifichi una situazione di carico eccessivo relativo ad un server e si voglia riportare il server in questione in uno stato stabile.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni attraverso diverse Common Table Expression (CTE):

1. Crea un Result Set *ErogazioniAttiveServer* con tutte le erogazioni al momento attive sul server fornito in input.

2. Crea un Result Set *ServerConContenuto* che esegue un join tra tutte le erogazioni identificate al punto precedente e i server che possiedono il contenuto erogato da ciascun erogazione, escluso quello corrente.
3. Crea un Result Set *CostoStreaming* che a partire da *ServerConContenuto* valuta il costo di streaming relativo l’utente che iniziato l’erogazione e ciascun server che possiede il contenuto in questione ed ha abbastanza banda per erogarlo.
4. Il Result Set finale identifica per ogni erogazione, attraverso il Result Set *CostoStreaming*, i server che presentano il costo di streaming minore per l’erogazione stessa.

I server forniti come output dalla query risultano dunque essere i server che, al momento della chiamata della funzione, presentano costo di streaming minore per il contenuto relativo all’erogazione in corso rispettivamente alla connessione dell’utente.

12.2.2 Gestione Excessive Memory

GestioneExcessiveMemory è una procedura che richiede come input l’id del server e propone l’eliminazione di determinati contenuti dalla cache del server. I contenuti proposti dovranno non essere in streaming su tale server al momento della chiamata e saranno mostrati in ordine crescente rispetto al numero di erogazioni, fornendo anche l’effettivo valore numerico, in modo da poterne valutare l’eliminazione dalla cache.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni:

1. Crea un Result Set *ErogazioniServer* che possiede tutte le erogazioni effettuate dal server fornito in input.
2. Crea un Result Set *NumeroErogazioniPerContenuto* che esegue un left join tra *Archiviazione* e il Result Set *ErogazioniServer* per valutare il totale di erogazioni per ogni contenuto posseduto dal server. All’interno del Result Set sono presenti solamente i contenuti che non sono erogati al momento della chiamata della funzione, in modo da poterne permettere l’eliminazione istantanea.
3. Il Result Set finale contiene per ogni contenuto in archivio sul server e non attualmente in streaming il numero totale di erogazioni in ordine crescente, ricavandole dal Result Set precedente *NumeroErogazioniPerContenuto*

Il risultato fornito dalla procedura permette di identificare rapidamente contenuti poco visualizzati che possono essere eliminati per riportare la memoria del server a valori più bassi, e permettere dunque l’immagazzinamento di nuovi contenuti.

12.2.3 Gestione Bilanciamento Carico

GestioneBilanciamentoCarico è una procedura che richiede come input l'id del server e suggerisce spostamenti di contenuti da un server per non sovraccaricarlo, indicando per ogni contenuto il server di destinazione. La chiamata alla procedura mostrerà per ogni contenuto presente in archivio sul server, da quello più erogato a quello meno erogato, il server più vicino al server dato come input che non ha già il contenuto archiviato. Gli spostamenti proposti da questa funzionalità avranno l'obiettivo di distribuire in maniera più omogenea i collegamenti ai server ed evitare sovraccarichi nei server.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni:

1. Crea un Result Set *ErogazioniServer* con tutte le erogazioni sul server fornito in input.
2. Crea un Result Set *NumeroErogazioniPerContenuto* che esegue un left join tra *Archiviazione* e il Result Set *ErogazioniServer* per valutare il totale di erogazioni per ogni contenuto posseduto dal server.
3. Crea un Result Set *DistanzeServer* che calcola la distanza tra il server in input e tutti gli altri server.
4. Il Result Set finale effettua un cross join tra *NumeroErogazioniPerContenuto* e *DistanzeServer* dando come risultato il prodotto cartesiano tra ogni contenuto erogato e ciascun server diverso da quello dato come input. Da questo risultato intermedio vengono selezionati soltanto i server che non hanno il contenuto in archivio e tra questi viene scelto quello con distanza minore dal server dato come input.

Il risultato finale permette di identificare i contenuti più erogati dal server e per ciascuno di questi il server dove è più conveniente spostare il contenuto, in modo da poter suddividere le erogazioni future in modo omogeneo.

12.3 Riprendi Visualizzazione

La funzionalità *Riprendi Visualizzazione* è progettata per migliorare l'esperienza utente su FilmSphere. Consente agli utenti di riprendere la visione di un film da dove l'hanno interrotto in modo semplice e intuitivo. Quando un utente accede alla piattaforma e seleziona questa opzione, il sistema effettua una query al database per identificare i film che sono stati iniziati ma non completati.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni:

1. Crea un Result Set *UltimeErogazioniUtente* con tutte le ultime erogazioni terminate dell'utente per ciascun film visualizzato almeno una volta.
2. Il Result Set finale per ciascun film mostra il titolo, l'id, i minuti visti e i minuti rimanenti per la visualizzazione dei film non visualizzati al completo.

Una volta che i dati sono stati recuperati con successo, l'utente riceve un elenco dei film non terminati insieme ai minuti rimasti e i minuti già visti. Questa informazione è estremamente utile perché consente agli utenti di prendere decisioni oculate su cosa riprendere a guardare in base al tempo disponibile e al punto esatto in cui hanno interrotto la visione.

Inoltre, questa funzionalità migliora l'usabilità della piattaforma, eliminando la necessità di ricordare manualmente dove si è interrotta la visione di un film o cercare manualmente il punto esatto. In questo modo, FilmSphere rende l'esperienza di streaming video più comoda e user-friendly per gli utenti, incoraggiando la continuazione della visione dei contenuti e l'ottimizzazione del tempo di fruizione.

12.4 Statistiche Lingue

La funzionalità Statistiche Lingue è stata progettata per fornire un'analisi dettagliata della distribuzione delle visualizzazioni nelle diverse lingue disponibili e dei sottotitoli su ogni film della piattaforma. Questo strumento offre una panoramica completa delle preferenze linguistiche degli utenti e consente di comprendere meglio la diversità linguistica della comunità di FilmSphere.

La procedura richiede come input una data di inizio e una data di fine del periodo di interesse su cui analizzare le statistiche relative alle lingue utilizzate per ogni film.

A livello implementativo la funzionalità viene gestita attraverso una query che svolge le seguenti operazioni:

1. Crea un Result Set *ErogazioniTotali* con tutte le erogazioni presenti nel database della piattaforma che fanno parte dell'orizzonte temporale indicato in input.
2. Crea un Result Set *VisualizzazioniTotaliFilm* con il numero totale di visualizzazioni per ogni film.
3. Crea un Result Set *ErogazioniPerSottitoliEAudio* con il numero totale di visualizzazioni per ogni film visualizzato in un determinata lingua audio e con associati i sottotitoli visti insieme, se presenti.
4. Il Result Set finale per ciascun film mostra l'id, il titolo, il nome della lingua dell'audio e dei sottotitoli, il numero visualizzazioni e la percentuale con cui è stato visualizzato con quella specifica coppia audio-sottotitoli rispetto al totale di visualizzazioni del film.

È possibile visualizzare le informazioni attraverso la procedura *StatisticheAudioSottotitoli* che permette anche di filtrare le erogazioni considerate per data utilizzando i parametri *dataInizio* e *dataFine*.

Queste informazioni sono preziose per la piattaforma poiché aiutano a comprendere meglio le esigenze linguistiche dei nostri utenti e a ottimizzare l'offerta di contenuti e servizi. Ad esempio, se risulta che un film è ampiamente visualizzato in una lingua specifica, potrà essere deciso di aggiungere più opzioni di sottotitoli in quella lingua o di offrire doppiaggi aggiuntivi.

Inoltre, queste statistiche percentuali possono essere utilizzate per orientare le strategie di marketing e promozione dei film in base alle preferenze linguistiche delle diverse audience, migliorando così l'esperienza globale degli utenti sulla piattaforma di streaming video FilmSphere.