



UNIVERSITÀ DI PISA

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

# Progetto del Corso di Basi di Dati

*Lorenzo Leoncini, Giulio Zingrillo*

ANNO ACCADEMICO 2022-2023

# Sommario

1 —	Introduzione . . . . .	1
2 —	Progettazione concettuale . . . . .	2
	2.1 Dizionario delle entità . . . . .	2
	2.2 Dizionario delle relationship . . . . .	8
	2.3 Legenda . . . . .	14
3 —	Ristrutturazione . . . . .	14
	3.1 Eliminazione delle generalizzazioni . . . . .	14
	3.2 Eliminazione degli attributi multivalore . . . . .	15
	3.3 Analisi ed eventuale eliminazione delle ridondanze . . . . .	15
	3.4 Partizionamenti o accorpamenti di entità e relationship . . . . .	15
4 —	Individuazione di operazioni sui dati . . . . .	16
	4.1 Rating assoluto, o Valutazione . . . . .	16
	4.2 Rating relativo, o Affinità . . . . .	22
	4.3 Raccomandazione di contenuti . . . . .	31
	4.4 Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione . . . . .	36
	4.5 Caching . . . . .	40
	4.6a Registrazione di un utente . . . . .	45
	4.6b Sottoscrizione del servizio da parte di un utente . . . . .	46
	4.7 Emissione di una fattura . . . . .	48
	4.8 Inserimento di una recensione . . . . .	50
	4.9 Analytics - Bilanciamento del carico . . . . .	51
	4.10a Custom Analytics - Fruizione media dei vincoli dell'abbonamento . . . . .	52
	4.10b Custom Analytics - Lingue per tempo di fruizione, come audio e come sottotitoli . . . . .	55
	4.11 Analytics - Classifiche . . . . .	56
	4.12 Operazione di fine erogazione . . . . .	59
5 —	Analisi delle prestazioni delle operazioni . . . . .	60
	5.1 Tavola dei volumi . . . . .	60
	5.2 Tavola delle operazioni . . . . .	66
	5.3 Analisi delle ridondanze . . . . .	68
6 —	Progettazione Logica . . . . .	97
	6.1 Descrizione dello schema logico . . . . .	98
	6.2 Implementazione . . . . .	98
	6.3 Vincoli . . . . .	107
7 —	Analisi delle dipendenze funzionali e normalizzazione . . . . .	110
8 —	Data Analytics . . . . .	111
	8.1 Classifiche . . . . .	112
	8.2 Bilanciamento del carico . . . . .	112
	8.3 Fruizione media dei vincoli dell'abbonamento . . . . .	113
	8.4 Lingue per tempo di fruizione, come audio e come sottotitoli . . . . .	114
9 —	Popolamento del database . . . . .	114

# 1 — Introduzione

La traccia del Progetto propone di realizzare una Base di Dati per l'implementazione di un servizio di memorizzazione e streaming di contenuti video online. Il portale memorizza anche informazioni sensibili degli utenti, come dati di fatturazione, abbonamenti e connessioni, e informazioni tecniche, relative, ad esempio, alla codifica dei film, alla dislocazione dei contenuti nei server che compongono la CDN, alle connessioni al servizio.

In fase di progettazione abbiamo fatto alcune ipotesi di lavoro, che riteniamo rendano la nostra trattazione più realistica.

La prima concerne il formato, che nelle *specifiche* assume una duplice accezione:

- Modalità di compressione dell'audio o del video di un film (e.g. AVI, o MP3), in questo documento denominata **formato**;
- Porzione dell'audio o del video di un film codificata in un formato, caratterizzata da una dimensione in memoria e da una durata che, come si evince dalle *specifiche*, non coincide necessariamente con la durata del film. Nel seguito chiameremo questo concetto **contenuto** o **file**. La codifica di un film in un formato può ripartirsi tra più contenuti, e.g. uno per la prima ora e uno per il resto del film. Inoltre, la stessa scena, in due contenuti dello stesso formato, può variare, ad esempio censurando contenuti sensibili, e per questo la visione di alcuni contenuti può essere interdetta in determinate aree geografiche<sup>1</sup>.

La seconda concerne i c.d. *contenuti esclusivi* e il loro legame con gli abbonamenti. Giacchè dalle *specifiche* appare evidente che lo stesso film può essere visualizzato con diverse definizioni in diversi piani di abbonamento, l'incidenza dell'abbonamento sul materiale visualizzabile dall'utente non può dispiegarsi sul piano del film, ma solo su quello del contenuto. Se, quindi, la maggior parte dei film è visualizzabile con tutti i piani di abbonamento - solo, con contenuti di qualità inferiore - esistono film che possono essere visti solo con piani di abbonamento specifici, e tali film sono, in ultima analisi, i *contenuti esclusivi*. In altre parole, un film *f* è un *contenuto esclusivo* se esiste un piano di abbonamento in cui *f* non è visualizzabile in nessun modo.

Nelle varie sezioni del progetto ci siamo poi trovati a fare una lunga serie di considerazioni e assunzioni, che sono troppo specifiche per essere riportate qui. Sarà nostra cura discuterle caso per caso, nelle note a piè di pagina e nelle tabelle, quando non direttamente nel corpo dei paragrafi.

## 2 — Progettazione concettuale

### 2.1 Dizionario delle entità

Gli attributi e le entità indicati con l'asterisco (\*) sono stati eliminati in fase di ristrutturazione; gli attributi e le entità indicati con due asterischi (\*\*) sono state aggiunti in fase di ristrutturazione; gli attributi indicati con tre asterischi (\*\*\*) sono ridondanti; gli attributi indicati tra parentesi quadre sono stati eliminati durante la normalizzazione (si rimanda alla SEZIONE 7). Gli attributi dei quali non è indicata la cardinalità sono da intendersi (1, 1). Gli attributi composti sono stati accorpati, in fase di ristrutturazione, sulle rispettive entità. Si veda 3.4.1 GESTIONE DEGLI ATTRIBUTI COMPOSTI.

Entità	Attributi	Identificatori	Descrizione
--------	-----------	----------------	-------------

---

<sup>1</sup>Come suggerito, per le restrizioni geografiche relative alla visualizzazione di contenuti, anche sulla base del piano di abbonamento, si è adottata un livello di granularità coincidente con lo Stato

Abbonamento	Nome, Tariffa, Durata (0, 1), MaxOre (0, 1), EtàMinima (0, 1), MaxGB (0, 1), Funzionalità Aggiuntiva* (0, N)	Nome	I vari piani di abbonamento in cui si articola l'offerta di FilmSphere. La <i>Tariffa</i> si intende misurata in euro mensili; la <i>Durata</i> si intende misurata in mesi: nel caso in cui l'attributo sia assente, l'abbonamento si rinnova ogni mese salvo disdetta da parte dell'utente. Per sottoscrivere un abbonamento può essere necessaria un'età minima. L'abbonamento, inoltre, può porre dei limiti alla mole di dati scaricata e alle ore di intrattenimento fruite mensilmente.
Artista	Id, Nome, Cognome, Pseudonimo (0, 1), Profilo, Attore**, Regista**	Id	Una figura professionale che lavora a un film. FilmSphere memorizza solo informazioni di attori e registi. Abbiamo aggiunto l'attributo <i>Pseudonimo</i> per avvicinare il nostro modello concettuale alla realtà. Il <i>Profilo</i> è una breve introduzione all'artista.
Attore*	Deriva tutti i propri attributi da ARTISTA	È identificato tramite ARTISTA	Recita in un film e può ricevere premi per le sue interpretazioni. Le comparse non sono memorizzate come attori.
Carta di Credito	Numero, CVV, Nome dell'intestatario, Cognome dell'intestatario, Mese di scadenza, Anno di scadenza	Numero	È l'unico mezzo di pagamento su FilmSphere. L'intestatario, come è ovvio, non coincide necessariamente con l'utente che paga tramite la carta. Gli attributi <i>Intestatario</i> e <i>Scadenza</i> sono composti nello schema ER. Il circuito della carta non è riportato in quanto ricavabile univocamente dal numero. Per mantenere i dati delle fatture passate, le carte di credito scadute rimangono memorizzate nel sistema.

Connessione	IP, Inizio, Fine (0, 1)	Inizio, identificazione esterna tramite la relationship con Dispositivo	Gli attributi Inizio e Fine memorizzano dei timestamp relativi alla connessione. Facendo riferimento a dispositivi per uso domestico, ipotizziamo che un dispositivo non possa stabilire nello stesso istante due diverse connessioni a FilmSphere, benché possa mantenere contemporaneamente più connessioni con esso.
Contenuto	Id, Dimensione, Lunghezza	Id	La <i>Dimensione</i> si intende in byte; la <i>Lunghezza</i> in secondi. Per la definizione di contenuto si rimanda all'INTRODUZIONE.
Critico	Id, Nome, Cognome, Pseudonimo (0, 1)	Id	Figura distinta dall'utente, il critico è un professionista abilitato a recensire film. <sup>2</sup> Le recensioni dei critici sono tradotte in tutte le lingue direttamente da FilmSphere. Come per gli artisti, abbiamo ritenuto opportuno inserire un attributo opzionale <i>Pseudonimo</i> .
Dispositivo	Id, Nome, Versione App, Sistema Operativo	Id	Con <i>Nome</i> si intende il nome commerciale del dispositivo; a FilmSphere è possibile accedere tramite applicazione ufficiale o tramite browser; in <i>Versione App</i> è registrata la versione dell'applicazione o del browser. Più utenti possono essere connessi a FilmSphere dallo stesso dispositivo.
Erogazione	Id, Inizio, Fine (0, 1)	Id	La fornitura di un contenuto a una connessione da parte di un server. Gli attributi <i>Inizio</i> e <i>Fine</i> sono dei timestamp.

---

<sup>2</sup>Per la differenza ontologica dei ruoli, abbiamo preferito non inserirlo nella generalizzazione di ARTISTA, nonostante abbia i medesimi attributi. Inoltre, nonostante sia accomunato all'UTENTE per la possibilità di fare recensioni, abbiamo preferito non creare una generalizzazione tra i due per la mole di attributi e relationship che afferiscono all'utente, la quale avrebbe reso tale generalizzazione fuorviante e pesantemente asimmetrica.

Fattore	Nome	Nome	Fondamentali per la profilazione dell'utente, i fattori indicano i parametri sulla base dei quali è possibile prevedere l'affinità di un utente con un film. Si faccia riferimento a 4.2 RATING RELATIVO, o AFFINITÀ I fattori vengono indicati, con il relativo <i>Valore</i> , dall'utente in fase di registrazione.
Fattura	Id, Saldo (0, 1), [Scadenza], Emissione	Id	<i>Saldo</i> , <i>Scadenza</i> e <i>Emissione</i> sono dei timestamp. L'attributo <i>Saldo</i> è opzionale: una fattura può non essere ancora stata pagata.
Film	Id, Titolo, Descrizione, Anno, Durata, Genere*(1, N), Somma Critica***, Totale Critica***, Somma Utenti***, Totale Utenti***, Rating Assoluto***, Visualizzazioni***	Id	Gli audiovisivi, memorizzati nei CONTENUTI, che FilmSphere possiede. L' <i>Anno</i> designa l'anno di produzione, che supponiamo, come da <i>specifiche</i> , uguale all'anno di uscita.
Formato*	Codice, Famiglia, DataRilascio, Qualità, Bitrate	Codice	Per la definizione di <i>Formato</i> , si rimanda all'INTRODUZIONE. Esempi di <i>Famiglia</i> sono MKV o AVI. Formati diversi, perché rilasciati in momenti successivi o aventi specifiche tecniche diverse, possono appartenere alla stessa famiglia. È il caso, ad esempio, di versioni diverse di MKV. La <i>DataRilascio</i> è memorizzata, come è ovvio, con un tipo <i>Date</i> . La <i>Qualità</i> è implementata con una stringa, che può assumere tre valori: <b>High</b> , <b>Medium</b> e <b>Low</b> . Il <i>Bitrate</i> è misurato in byte per secondo. <i>Qualità</i> e <i>Bitrate</i> sono accorpate nell'attributo composto <i>Specifiche tecniche</i> . La generalizzazione è totale ed esclusiva.
Formato Audio	Deriva tutti i suoi attributi da FORMATO	È identificato tramite FORMATO	Costituisce il formato di codifica audio di un contenuto.

Formato Video	Attributi derivati da FORMATO, Risoluzione, Rapporto d'Aspetto	È identificato tramite FORMATO	Costituisce il formato di codifica video di un contenuto. La <i>Risoluzione</i> è identificata tramite un numero intero (e.g. 480, 720, 1080). Il <i>Rapporto d'aspetto</i> è pensato come una stringa, composta da due numeri separati da uno spazio.
Funzionalità Aggiuntiva**	Nome	Nome	Rappresenta una feature di FilmSphere che è fruibile solo tramite determinati piani di abbonamento. L'esempio fatto nelle <i>specifiche</i> concerne la riproduzione di film offline.
Genere**	Nome	Nome	Il genere del film, essenziale per le funzioni dell'area analytics e per le raccomandazioni.
Lingua	Nome	Nome	L'idioma in cui può essere doppiata, o sottotitolata, una versione del film.
Paese	Nome, InizioIP, FineIP, Latitudine, Longitudine	Nome	L'unica entità a non appartenere ad alcuna area tematica nel nostro schema Entity-Relationship, per la sua natura marcatamente trasversale. Rappresenta una nazione, la posizione del suo centro geometrico e l'intervallo di indirizzi IP ad essa associato. <sup>3</sup> Gli attributi <i>InizioIP</i> , <i>FineIP</i> , <i>Latitudine</i> , <i>Longitudine</i> sono accorpati in due attributi composti.

---

<sup>3</sup>Supponiamo che gli indirizzi IP siano tutti v4 e che gli intervalli di IP associati a ciascun Paese siano contigui, senza distinzioni tra reti di classe A, B e C. Alla luce di questa ipotesi se, ad esempio, gli indirizzi 1.1.1.1 e 1.1.1.3 sono italiani, allora anche l'indirizzo 1.1.1.2 è italiano.

Premio	Id, Nome, Anno, Istituzione, Descrizione, Peso, Attore**, Film**, Regista**	Id	Identifica l'istanza di premio. L' <i>Anno</i> è l'anno di attribuzione, non correlato con l'anno di uscita del film. L' <i>Istituzione</i> è l'organizzazione, pubblica o privata, che conferisce il premio. Il <i>Peso</i> , espresso in centesimi, indica la rilevanza del premio nelle metriche di rating. La generalizzazione non è esclusiva: alcuni premi possono lodare l'interpretazione di un attore in un film, e pertanto riferirsi contemporaneamente a entrambi. È però totale dal momento che FilmSphere non memorizza premi che non concernano film, attori o registi. <sup>4</sup>
Premio Attore*	Deriva tutti i suoi attributi da PREMIO	È identificato tramite PREMIO	Un premio attribuito a un attore.
Premio Film*	Deriva tutti i suoi attributi da PREMIO	È identificato tramite PREMIO	Un premio attribuito a un film.
Premio Regista*	Deriva tutti i suoi attributi da PREMIO	È identificato tramite PREMIO	Un premio attribuito a un regista.
Regista*	Deriva tutti i suoi attributi da ARTISTA	È identificato tramite ARTISTA	Dirige film e può, per il suo lavoro, ricevere premi.
Server	Id, LarghezzaBanda, CapacitàMax, Longitudine, Latitudine, Jitter, Banda Disponibile***	Id	I nodi della CDN, fondamentali per l'erogazione del servizio. La <i>Larghezza di banda</i> è misurata in byte per secondo; la <i>Capacità Massima</i> è misurata in byte. Gli attributi <i>Latitudine</i> e <i>Longitudine</i> sono uniti nell'attributo composto <i>Posizione</i> . L'attributo <i>Jitter</i> è un numero naturale tra 1 e 10 che misura la variabilità nella latenza della trasmissione del server: come esposto nel PARAGRAFO 4.4, un jitter elevato pregiudica notevolmente la qualità del servizio di streaming.

<sup>4</sup>Abbiamo scelto di modellizzare solo le istanze di premio, senza curarci di ulteriori generalizzazioni, per mantenere la nostra trattazione massimamente flessibile, dal momento che i vari premi possono seguire logiche molto eterogenee.



Utente	Codice, Cognome, Password, Nascita	Nome, E-mail, Data	Codice	Il fruitore dei contenuti di Film-Sphere. La password viene memorizzata cifrata tramite AES-256. Abbiamo ritenuto opportuno memorizzare la data di nascita, per gestire le restrizioni di età in fase di sottoscrizione degli abbonamenti e per prevedere i contenuti richiesti dall'utente in fase di caching. L'indirizzo e-mail identifica univocamente l'utente: in fase di accesso al portale, infatti, esso è l'unico modo che l'utente ha di identificare se stesso, dal momento che non conosce il proprio CODICE. <sup>5</sup>
--------	---	--------------------------	--------	---

## 2.2 Dizionario delle relationship

Le relationship indicate con l'asterisco (\*) sono state eliminate in fase di ristrutturazione; le relationship indicate con due asterischi (\*\*) sono state aggiunte in fase di ristrutturazione.

Relationship	Entità 1	Entità 2	Attributi
Accesso	Erogazione (1, 1) Ogni erogazione fa riferimento a uno e un solo contenuto.	Contenuto (0, N) Ogni contenuto può essere erogato più volte, oppure nessuna: è il caso, ad esempio, di un contenuto appena inserito nel database.	-
Appartenenza**	Genere (1, N) A ogni genere possono corrispondere uno o più film.	Film (1, N) Ogni film è riconducibile ad almeno un genere.	-
Audio	Lingua (0, N) Ogni lingua può essere l'audio di più film oppure di nessuno: è il caso, ad esempio, di una lingua appena inserita nel database.	Contenuto (0, 1) Ogni contenuto può essere relativo a una lingua, oppure a nessuna, nel caso in cui sia un contenuto video.	-

<sup>5</sup>Nonostante l'attributo *E-mail* sia chiave, non è chiave primaria: è infatti codificato come una stringa di grandi dimensioni, che renderebbe poco efficienti le operazioni sul database. Risulta pertanto preferibile aggiungere l'attributo *Codice*, come peraltro suggerito nelle *specifiche*.

Codifica*	<p>Formato (0, N)</p> <p>In un formato possono essere codificati più contenuti, oppure nessuno: è il caso, ad esempio, di un formato ormai obsoleto, ma supportato da FilmSphere.</p>	<p>Contenuto (1, 1)</p> <p>Ogni contenuto è codificato in un solo formato, che deve essere alternativamente Audio o Video.</p>	-
Codifica Audio**	<p>Formato Audio (0, N)</p> <p>In un formato possono essere codificati più contenuti, oppure nessuno: è il caso, ad esempio, di un formato ormai obsoleto, ma supportato da FilmSphere.</p>	<p>Contenuto (0, 1)</p> <p>Ogni contenuto è codificato in un solo formato, che deve essere alternativamente Audio o Video.</p>	-
Codifica Video**	<p>Formato Video (0, N)</p> <p>In un formato possono essere codificati più contenuti, oppure nessuno: è il caso, ad esempio, di un formato obsoleto.</p>	<p>Contenuto (0, 1)</p> <p>Ogni contenuto è codificato in un solo formato, che deve essere alternativamente Audio o Video.</p>	-
Collegamento	<p>Server (0, N)</p> <p>Ogni server può aver portato a termine più erogazioni, ma anche nessuna: è il caso, ad esempio, di un server appena inserito nella CDN.</p>	<p>Erogazione (1, 1)</p> <p>Ogni erogazione fa riferimento a uno e un solo server.</p>	-
Direzione	<p>Film (1, N)</p> <p>Ogni film ha almeno un regista, ma può averne più di uno (e.g. i fratelli Coen).</p>	<p>Regista, Artista dopo la ristrutturazione (0, N)</p> <p>Ogni regista può dirigere più film, ma anche nessuno, ad esempio nel caso in cui sia stato appena inserito nel database.</p>	-
Fruizione	<p>Connessione (0, N)</p> <p>Ogni connessione può fruire di più erogazioni, ma anche di nessuna: è il caso, ad esempio, di una connessione appena instaurata.</p>	<p>Erogazione (1, 1)</p> <p>Ogni erogazione fa riferimento a una e una sola connessione.</p>	-

Importanza	<p>Utente (0, N)</p> <p>Se un utente <math>u</math> e un fattore <math>f</math> non sono posti in relazione da questa relationship, allora per <math>u</math> <math>f</math> ha un peso pari a 0. Le preferenze di un utente possono essere influenzate da più fattori, oppure da nessuno.</p>	<p>Fattore (0, N)</p> <p>Un fattore può essere tenuto in considerazione da più utenti, oppure, potenzialmente, da nessuno.</p>	<p>Valore: indica, in centesimi, qual è il peso del fattore per l'utente. Viene indicato dall'utente stesso in fase di registrazione.</p>
Instaurazione	<p>Utente (1, N)</p> <p>Ogni utente può instaurare più connessioni con FilmSphere. Ne instaura sempre almeno una: quella con cui si è registrato nel database.</p>	<p>Connessione (1, 1)</p> <p>Ogni connessione fa riferimento a uno e un solo utente.</p>	-
Interpretazione	<p>Attore, Artista dopo la ristrutturazione (0, N)</p> <p>Un attore può aver interpretato diversi film, oppure nessuno: questo permetterebbe di conservare i dati di un attore, e la profilazione degli utenti a esso associata, anche nel caso in cui i film che interpreta non fossero più disponibili in piattaforma. Per i nostri scopi, il ruolo che l'attore interpreta nel film non rileva, e pertanto un attore che interpreta più ruoli nello stesso film è modellizzato come una interpretazione.</p>	<p>Film (0, N)</p> <p>Un film può essere interpretato da diversi attori oppure da nessuno: si pensi a un cartone animato senza dialoghi, come <i>Shawn the Sheep</i>.</p>	-
Intestazione	<p>Fattura (1, 1)</p> <p>Ogni fattura si riferisce a uno e un solo utente.</p>	<p>Utente (0, N)</p> <p>A ogni utente possono essere intestate più fatture, ma anche nessuna: si pensi, ad esempio, al caso di un utente appena registrato in piattaforma.</p>	-

Nazionalità	<p>Utente (1, 1)</p> <p>Ogni utente, su FilmSphere, dichiara una e una sola nazionalità.</p>	<p>Paese (0, N)</p> <p>Da un Paese possono provenire più utenti registrati, oppure nessuno. In questo caso, il Paese potrebbe essere egualmente presente nel database, ad esempio, per una restrizione geografica.</p>	-
Offerta Contenuto	<p>Abbonamento (1, N)</p> <p>Ogni piano d'abbonamento deve offrire almeno un contenuto, ma può offrirne più d'uno.</p>	<p>Contenuto (0, N)</p> <p>Ogni contenuto può essere offerto da più piani d'abbonamento, oppure da nessuno: è il caso, ad esempio, di un contenuto che non è ancora stato lanciato sulla piattaforma, ma che è già stato caricato sul database e sui server.</p>	-
Offerta Funzionalità**	<p>Abbonamento (0, N)</p> <p>Un abbonamento può offrire più funzionalità aggiuntive, oppure nessuna.</p>	<p>Funzionalità Aggiuntiva (0, N)</p> <p>Una funzionalità aggiuntiva può essere offerta da più piani di abbonamento, oppure da nessuno: è il caso, ad esempio, di una funzionalità implementata ma non ancora aperta agli utenti.</p>	-
Pagamento	<p>Carta di Credito (0, N)</p> <p>Una carta di credito può aver saldato più fatture, oppure nessuna.</p>	<p>Fattura (1, 1)</p> <p>Ogni fattura può essere saldata da una e una sola carta di credito.</p>	-
Possesso Server	<p>Contenuto (1, N)</p> <p>Ogni contenuto può essere posseduto da più server, ma deve essere posseduto da almeno uno, altrimenti non è posseduto da FilmSphere.</p>	<p>Server (0, N)</p> <p>Ogni server può possedere più contenuti, ma può anche non possederne alcuno: è il caso, ad esempio, di un server appena aggiunto alla CDN.</p>	-

Possesso Utente	<p>Utente (0, 1)</p> <p>Ogni utente può essere associato a una sola carta di credito, che sarà utilizzata per generare le fatture. La creazione del profilo, come riportato ai PARAGRAFI 4.6 E 4.7, è divisa in due fasi: la prima prevede la registrazione dell'utente; la seconda l'associazione di una carta di credito e la scelta di un piano di abbonamento. Possono pertanto esistere utenti a cui non è associata alcuna carta di credito.</p>	<p>Carta di Credito (1, N)</p> <p>Ogni carta di credito è associata a un utente, ma potrebbe essere associata a più d'uno.</p>	-
Premiazione Attore	<p>Attore, Artista dopo la ristrutturazione (0, N)</p> <p>Ogni attore essere insignito di più premi, oppure di nessuno.</p>	<p>Premio Attore, Premio dopo la ristrutturazione (1, N)</p> <p>Per mantenere la massima flessibilità, ogni premio può essere assegnato a più attori. FilmSphere memorizza un premio solo se esso è stato assegnato a un attore memorizzato nel database.</p>	-
Premiazione Film	<p>Film (0, N)</p> <p>Ogni film può essere insignito di più premi, oppure di nessuno.</p>	<p>Premio Film, Premio dopo la ristrutturazione (1, N)</p> <p>Per mantenere la massima flessibilità, ogni premio può essere assegnato a più film. FilmSphere memorizza un premio solo se esso è stato assegnato a un film memorizzato nel database.</p>	-
Premiazione Regista	<p>Regista, Artista dopo la ristrutturazione (0, N)</p> <p>Ogni regista può essere insignito di più premi, oppure di nessuno.</p>	<p>Premio Regista, Premio dopo la ristrutturazione (0, N)</p> <p>Per mantenere la massima flessibilità, ogni premio può essere assegnato a più registi. FilmSphere memorizza un premio solo se esso è stato assegnato a un regista memorizzato nel database.</p>	-
Produzione	<p>Film (1, 1)</p> <p>Ogni film viene prodotto in uno e un solo Paese.</p>	<p>Paese (0, N)</p> <p>In un Paese possono essere stati prodotti più film, oppure nessuno.</p>	-

Rappresentazione	Film (1, N) Ogni film è rappresentato in uno o più contenuti.	Contenuto (1, 1) Ogni contenuto rappresenta uno e un solo film.	-
Recensione Critico	Critico (0, N) Ogni critico può essere autore di più recensioni, oppure di nessuna: è il caso, ad esempio, di un critico il cui profilo è appena stato inserito nel database.	Film (0, N) Ogni film può essere stato recensito da più critici, oppure da nessuno.	Data, Voto, Testo. Il voto è espresso in quinti.
Recensione Utente	Utente (0, N) Ogni utente può essere autore di più recensioni, oppure di nessuna.	Film (0, N) Ogni film può essere stato recensito da più utenti, oppure da nessuno	Data, Voto, Testo. Il voto è espresso in quinti.
Restrizione Abbonamento <sup>6</sup>	Abbonamento (0, N) Un abbonamento può prevedere restrizioni in più Paesi, oppure in nessuno.	Paese (0, N) Un Paese può essere oggetto di restrizioni da parte di più abbonamenti, oppure di nessuno.	-
Restrizione Contenuto	Contenuto (0, N) La visione di un contenuto può essere interdetta in più Paesi, oppure in nessuno.	Paese (0, N) In un Paese può essere interdetta la visione di più contenuti, oppure di nessuno.	-
Servizio	Abbonamento (0, N) A un abbonamento possono essere relative più fatture, oppure nessuna.	Fattura (1, 1) Ogni fattura è relativa a uno e un solo piano di abbonamento. Questa relationship è necessaria dal momento che FilmSphere conserva i dati relativi a fatture passate.	-
Sottoscrizione	Abbonamento (0, N) Un piano di abbonamento può essere sottoscritto da più utenti, oppure da nessuno.	Utente (0, 1) Un utente non può avere due piani di abbonamento contemporaneamente. Può, però, non averne alcuno, ad esempio se quello che aveva è scaduto o se l'utente è appena stato registrato nel database (per questo secondo caso si vedano i paragrafi 4.6 e 4.7).	Inizio: la data in cui l'utente ha sottoscritto l'abbonamento corrente.

<sup>6</sup>Un abbonamento si considera oggetto di restrizione in un Paese se non può essere sottoscritto da un abitante di quel Paese.  
Cfr. SEZIONE 9 — POPOLAMENTO DEL DATABASE

Sottotitoli	<p>Contenuto (0, N)</p> <p>Un contenuto può essere sottotitolato in più lingue, oppure in nessuna. Assumiamo in questa sede che i file con i sottotitoli nelle varie lingue siano parte integrante di un contenuto video.</p>	<p>Lingua (0, N)</p> <p>I sottotitoli di più film possono essere scritti in una stessa lingua, ma è anche possibile che vi siano lingue in cui non sono scritti i sottotitoli di nessun film.</p>	-
Utilizzo	<p>Connessione (1, 1)</p> <p>Ogni connessione è relativa a uno e un solo dispositivo.</p>	<p>Dispositivo (1, N)</p> <p>A ogni dispositivo possono corrispondere più connessioni. A ogni dispositivo ne corrisponde sempre almeno una: quella nella quale il dispositivo è stato registrato e inserito nel database.</p>	-

## 2.3 Legenda

Nello schema Entità-Relationship allegato al presente documento vengono evidenziate, tramite colori, le diverse aree del database. In particolare:

- Il colore **blu** fa riferimento all'**Area Contenuti**;
- Il colore **rosso** fa riferimento all'**Area Formati**;
- Il colore **verde** fa riferimento all'**Area Clienti**;
- Il colore **arancione** fa riferimento all'**Area Streaming**;
- Il colore **nero** fa riferimento a entità e relationship non riconducibili a nessun area in particolare.

Gli attributi ridondanti sono indicati tramite il colore **grigio**.

L'attributo eliminato in fase di normalizzazione è segnalato tramite il colore **viola**.

## 3 — Ristrutturazione

### 3.1 Eliminazione delle generalizzazioni

#### 3.1.1 Artista

L'entità ARTISTA è specializzata in ATTORE e REGISTA. La generalizzazione è totale e non esclusiva. Le entità figlie non hanno attributi propri. Possiamo pertanto accorpare i figli nel padre, introducendo gli attributi booleani *Regista* e *Attore* e i seguenti vincoli: per dirigere un film o ricevere un premio regista è necessario essere registi; per interpretare un film o ricevere un premio attore è necessario essere un attore.

#### 3.1.2 Formato

L'entità FORMATO è specializzata in FORMATO AUDIO e FORMATO VIDEO. La generalizzazione è totale ed esclusiva. Dal momento che non vi sono accessi all'entità padre che non siano anche accessi alle entità figlie, la traduzione della generalizzazione con una relationship è da escludersi. Poiché inoltre l'entità FORMATO VIDEO presenta due attributi propri è conveniente tradurre la generalizzazione accorpendo il genitore nei

figli e introducendo due relationship tra i figli e l'entità CONTENUTO. Bisogna introdurre il vincolo che ogni entità di contenuto deve partecipare a esattamente una relationship tra le due appena create.

### 3.1.3 Premio

L'entità PREMIO è specializzata in PREMIO ATTORE, PREMIO FILM e PREMIO REGISTA. La generalizzazione è totale e non esclusiva. Dal momento che non vi sono accessi all'entità padre che non siano anche accessi alle entità figlie, la traduzione della generalizzazione con un relationship è da escludersi. Poiché inoltre le entità figlie non hanno attributi propri possiamo accorpare i figli nel genitore introducendo gli attributi booleani *Attore*, *Film* e *Regista*, e i seguenti vincoli: un premio che sia attribuito a un regista deve essere un premio regista; un premio che sia attribuito a film deve essere un premio film; un premio che sia attribuito a un attore deve essere un premio attore.

## 3.2 Eliminazione degli attributi multivalore

### 3.2.1 Genere

L'attributo multivalore *Genere*, con cardinalità (1, N) viene sostituito dall'entità GENERE, con un unico attributo *Nome*, che la identifica, e dalla relationship APPARTENENZA, con cardinalità (1, N) verso GENERE e (1, N) verso FILM.

### 3.2.2 Funzionalità Aggiuntiva

L'attributo multivalore *Funzionalità Aggiuntiva*, con cardinalità (0, N) viene sostituito dall'entità FUNZIONALITÀ AGGIUNTIVA, con un unico attributo *Nome*, che la identifica, e dalla relationship OFFERTA FUNZIONALITÀ, con cardinalità (0, N) verso ABBONAMENTO e (0, N) verso FUNZIONALITÀ AGGIUNTIVA.

## 3.3 Analisi ed eventuale eliminazione delle ridondanze

Si rimanda al paragrafo 5: ANALISI DELLE PRESTAZIONI DELLE OPERAZIONI.

## 3.4 Partizionamenti o accorpamenti di entità e relationship

Non è stato necessario effettuare partizionamenti o accorpamenti.

### 3.4.1 Gestione degli attributi composti

Gli attributi composti *Posizione* (in relazione a PAESE e SERVER), *Intestatario*, *Scadenza*, *Specifiche tecniche* e *IP* rappresentano tutti proprietà la cui ripetizione in istanze distinte dell'entità a cui si riferiscono è rara, quando non proprio impossibile (si pensi al caso di PAESE). Si è pertanto scelto di accorpare ogni attributo composto sull'entità a cui si riferisce. Questo ha comportato alcune modifiche alla nomenclatura, riportate di seguito:

Attributo composto	Entità	Attributo originario	Attributo ristrutturato
Intestatario	Carta di Credito	Nome	Nome Intestatario
Intestatario	Carta di Credito	Cognome	Cognome Intestatario
IP	Paese	Inizio	Inizio IP
IP	Paese	Fine	Fine IP



Scadenza	Carta di Credito	Mese	Mese Scadenza
Scadenza	Carta di Credito	Anno	Anno Scadenza

I nomi di tutti gli altri attributi sono rimasti invariati.

## 4 — Individuazione di operazioni sui dati

Le frequenze delle operazioni e i volumi riportati in questa sezione sono il risultato delle stime fatte nella SEZIONE 5 — ANALISI DELLE PRESTAZIONI. In tale sezione, inoltre, alcune delle operazioni che seguono saranno rese più efficienti tramite l'introduzione di ridondanze. Gli accessi nelle tavole dedicate si intendono per ogni esecuzione dell'operazione.

Nella determinazione degli accessi abbiamo sempre mirato a rappresentare il caso peggiore, considerando come puntuali i soli accessi su chiave primaria e leggendo per intero la relazione in tutti gli altri casi. Questo approccio è a nostro avviso migliore di tanti altri nella misura in cui rimane avulso da riferimenti all'implementazione fisica della base di dati e non indulge a stime grossolane.<sup>7</sup>

### 4.1 Rating assoluto, o Valutazione

**Input:** Un film.

**Output:** Un numero razionale tra 0 e 5.

**Frequenza Stimata:** 160mila volte al giorno.

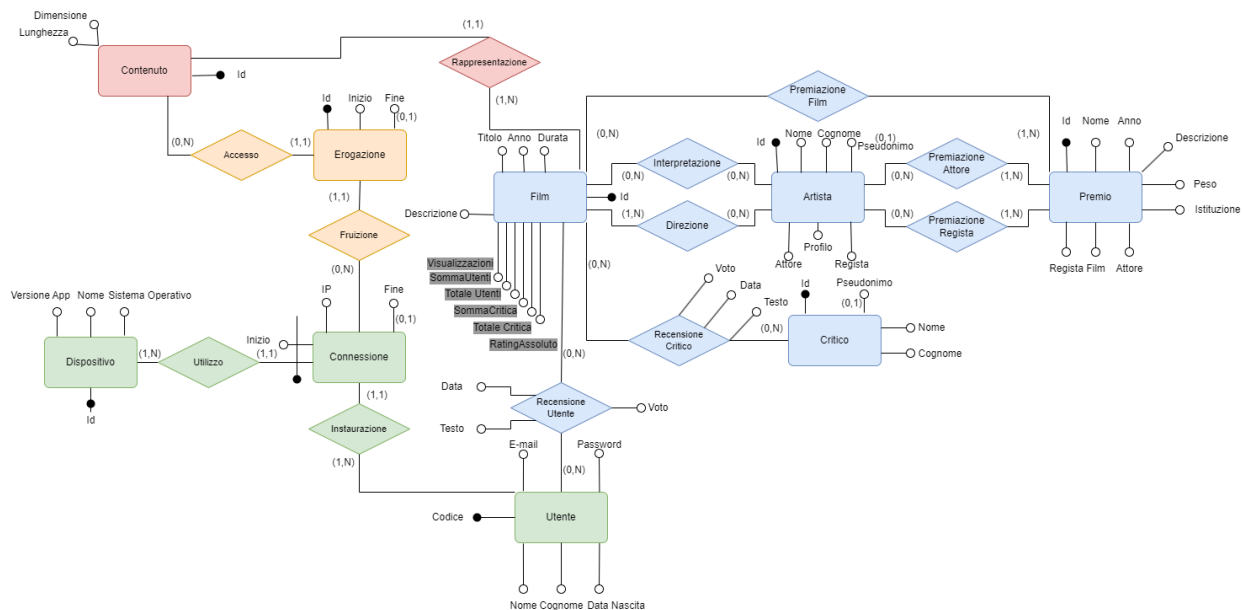


Figura 1: Diagramma ER coinvolto dall'operazione 1.

<sup>7</sup>Premettiamo una nota alle tavole degli accessi: in taluni casi si è preferito non dedicare troppo spazio a un'analisi, statistica e probabilistica, che avesse minime ripercussioni sui risultati finali. Nei casi in cui, invece, il divario numerico era rilevante abbiamo ricercato il massimo rigore, a costo di complicare notevolmente la trattazione.

Il rating assoluto, rappresentato in stelle, esprime una valutazione generale del film. Viene calcolato come media pesata dei seguenti fattori:

Valore	Peso
Media delle recensioni della critica	10 %
Media delle recensioni degli utenti	20 %
Fascia di premialità	10 %
Fascia di visualizzazioni del film <sup>8</sup>	40 %
Celebrità dei registi	10 %
Celebrità degli attori	10 %

#### 4.1.1 Recensioni

Le recensioni della critica e degli utenti sono attribuite in quinti; se ne fa la media aritmetica.

#### 4.1.2 Fascia di premialità

Si ricava sommando i pesi dei premi vinti dal film. Utilizzando la stima, fatta nella TAVOLA DEI VOLUMI, di 1.5 premi vinti da ogni film di FilmSphere, e ipotizzando che ogni premio abbia un peso medio di 50 possiamo elaborare la tabella seguente.

Fascia	Somma dei pesi
0	Minore di 30
1	Tra 30 e 50
2	Tra 50 e 70
3	Tra 70 e 90
4	Tra 90 e 110
5	Maggiore di 110

---

<sup>8</sup>Qui e nel RATING RELATIVO le visualizzazioni sono modellizzate come *erogazioni consecutive di contenuti rappresentanti lo stesso film*. In altre parole, la visualizzazione coincide con l'erogazione se e solo se il film rappresentato dall'ultimo contenuto erogato alla connessione è diverso da quello rappresentato dal contenuto attualmente in erogazione. Questa definizione di visualizzazione è, tra quelle che proponiamo, la più rigorosa e la più pesante da un punto di vista computazionale: in alcuni casi, perciò, le preferiremo altre metriche meno precise ma più efficienti.

#### 4.1.3 Fascia di visualizzazioni del film

La tabella che segue è stata realizzata utilizzando le seguenti stime, riportate nella TAVOLA DEI VOLUMI:

- Gli utenti attivi in un istante casuale del primo anno di funzionamento di FilmSphere sono stati mediamente 65mila<sup>9</sup>;
- Ogni utente attivo visualizza in media un film al giorno.

Ogni film è pertanto stato visualizzato in media  $365 \cdot 65000 / 500 = 47450$  volte dal lancio di FilmSphere.

Fascia	Numero di visualizzazioni
0	Minore di 20mila
1	Tra 20mila e 30mila
2	Tra 30mila e 40mila
3	Tra 40mila e 50mila
4	Tra 50mila e 60mila
5	Maggiore di 60mila

#### 4.1.4 Celebrità dei registi

È data dalla media aritmetica tra la fascia di premialità e la fascia di visualizzazioni del regista, calcolate come riportato di seguito. Nella TAVOLA DEI VOLUMI abbiamo stimato che ogni regista presente su FilmSphere abbia ricevuto mediamente, nel corso della sua carriera, 0.8 premi.

Fascia	Somma di pesi dei premi ricevuti dal regista
0	Minore di 20
1	Tra 20 e 30
2	Tra 30 e 40
3	Tra 40 e 50
4	Tra 50 e 60
5	Oltre 60

La seconda tabella è ricavata considerando le seguenti stime, provenienti dalla TAVOLA DEI VOLUMI e dal PARAGRAFO 4.1.3:

---

<sup>9</sup>Qui e di seguito, si definisce *attivo* un utente a cui sia associato un abbonamento in corso di validità.

- I registi presenti su FilmSphere sono 400;
- La relationship DIREZIONE ha volume pari a 600;
- Ogni film è stato visualizzato in media, dal lancio di FilmSphere, 47450 volte.

Ne segue che ogni regista ha diretto in media 1.5 film: basta pertanto moltiplicare tutti i valori della tabella del PARAGRAFO 4.1.3 per 1.5:

<b>Fascia</b>	<b>Numero di visualizzazioni di tutti i film diretti dal regista</b>
0	Minore di 30mila
1	Tra 30mila e 45mila
2	Tra 45mila e 60mila
3	Tra 60mila e 75mila
4	Tra 75mila e 90mila
5	Oltre 90mila

Nel caso in cui vi siano più registi, ai fini della determinazione della fascia si fa la media aritmetica tra le visualizzazioni e i pesi dei premi vinti da ogni regista.

#### 4.1.5 Celebrità degli attori

È data dalla media aritmetica tra la fascia di premialità e la fascia di visualizzazioni dell'attore, calcolate come riportato di seguito. Per la prima tabella abbiamo stimato il peso medio di un premio a 50 e utilizzato l'ipotesi, proveniente dalla TAVOLA DEI VOLUMI, che ogni attore presente su FilmSphere abbia ricevuto in media 0.4 premi

<b>Fascia</b>	<b>Somma dei pesi dei premi ricevuti dall'attore</b>
0	Minore di 5
1	Tra 5 e 10
2	Tra 10 e 20
3	Tra 20 e 30
4	Tra 30 e 40
5	Maggiore di 40

La seconda tabella è ottenuta moltiplicando tutti i valori della tabella al PARAGRAFO 4.1.3 per 2, alla luce della stima, fatta nella TAVOLA DEI VOLUMI che ogni attore presente su FilmSphere interpreti in media 2 film.

Fascia	Numero di visualizzazioni di tutti i film interpretati dall'attore
0	Minore di 40mila
1	Tra 40mila e 60mila
2	Tra 60mila e 80mila
3	Tra 80mila e 100mila
4	Tra 100mila e 120mila
5	Maggiore di 120mila

Nel caso in cui vi siano più attori, ai fini della determinazione della fascia si fa la media aritmetica tra le visualizzazioni e i pesi dei premi vinti da ogni attore. Ricordiamo che le comparse non sono memorizzate tra gli attori.

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Artista	E	2960
Connessione	E	47.45 milioni
Contenuto	E	28mila
Critico	E	30
Direzione	R	600
Erogazione	E	94.9 milioni
Film	E	500
Fruizione	R	94.9 milioni
Instaurazione	R	47.45 milioni
Interpretazione	R	5600

Premiazione Attore	R	1120
Premiazione Film	R	750
Premiazione Regista	R	320
Premio	E	2mila
Rappresentazione	R	28mila
Recensione Critico	R	2mila
Recensione Utente	R	300mila
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico	R	2mila	R	Leggiamo il voto delle recensioni della critica relative al film.
2	Recensione Utente	R	300mila	R	Leggiamo il voto delle recensioni degli utenti relative al film.
3	Premiazione Film	R	750	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	5600	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	600	R	Leggiamo i film diretti dai registi del film di nostro interesse.
6	Premiazione Attore	R	1120	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	320	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.

8	Premio	E	6.86	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.
9	Rappresentazione	R	28mila	R	Leggiamo i contenuti che rappresentano il film di nostro interesse e gli altri 11 film in cui hanno recitato gli attori.
10	Accesso	R	94.9 milioni	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto.
11	Fruizione	R	94.9 milioni	R	Per ogni erogazione, leggiamo la connessione destinataria.
12	Instaurazione	R	47.45 milioni	R	Per ogni connessione, leggiamo l'utente relativo. Questa azione è fondamentale per ricavare il numero di visualizzazioni.

## 4.2 Rating relativo, o Affinità

**Input:** Un film, un utente.

**Output:** Un numero razionale tra 0 e 100.

**Frequenza stimata:** 160mila volte al giorno.

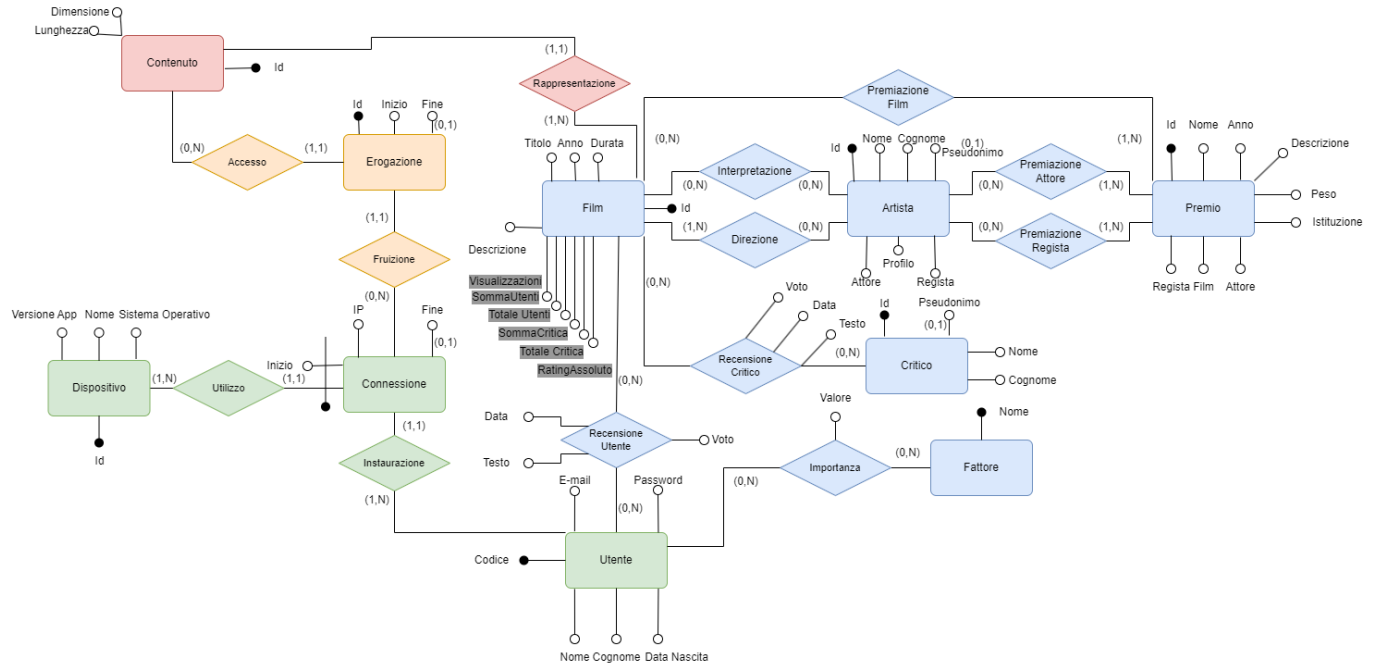


Figura 2: Diagramma ER coinvolto dall'operazione 2.

Il rating relativo quantifica l'affinità che un utente ha con un film. In fase di registrazione, l'utente seleziona dei fattori, e a ognuno di essi associa un valore in centesimi. Il rating relativo è la media, pesata sui valori, di tali fattori. Ne segue che attribuire un valore elevato a un fattore significa dargli grande importanza nella determinazione del proprio rating relativo. I fattori selezionabili dall'utente sono i seguenti:

- *Film d'autore*: film con registi celebri; la celebrità di un regista è calcolata con metriche più specifiche rispetto al RATING ASSOLUTO, e in centesimi;
- *I film che hanno fatto la storia*: il fattore è costituito dal numero di anni compresi tra la data di uscita del film e la data locale al database. In caso tale numero sia maggiore di 100, viene troncato a 100;
- *Esaltati dalla critica*: viene valutata la media delle recensioni della critica, convertita in centesimi;
- *I più amati*: viene valutata la media delle recensioni degli utenti, convertita in centesimi;
- *Popolari*: viene valutata la fascia di visualizzazioni del film, in centesimi;
- *Le star*: film con attori celebri; la celebrità degli attori è calcolata con metriche più specifiche rispetto al RATING ASSOLUTO, e in centesimi;
- *Premiati*: viene valutata la fascia di premialità del film, in centesimi;

Le tabelle che seguono sono una parziale riproposizione di quelle dell' OPERAZIONE 1, ma sono maggiormente dettagliate in considerazione della maggiore precisione richiesta dal calcolo.

#### 4.2.1 Celebrità dei registi

È data dalla media aritmetica tra la fascia di premialità e la fascia di visualizzazioni del regista, calcolate come riportato sotto<sup>10</sup>:

<b>Fascia</b>	<b>Somma dei pesi dei premi ricevuti dal regista</b>
---------------	--

<sup>10</sup>Qui e di seguito, essendo il numero di premi una grandezza discreta ad alta granularità, vi associamo fasce del 10%. Per le visualizzazioni, sempre discrete ma con granularità più fine, utilizzeremo invece fasce del 5%.



0	Minore di 20
10	Tra 20 e 25
20	Tra 25 e 30
30	Tra 30 e 35
40	Tra 35 e 40
50	Tra 40 e 45
60	Tra 45 e 50
70	Tra 50 e 55
80	Tra 55 e 60
90	tra 60 e 70
100	Maggiore di 70

<b>Fascia</b>	<b>Numero di visualizzazioni di tutti i film diretti dal regista</b>
0	Minore di 30000
5	Tra 30000 e 33750
10	Tra 33750 e 37500
15	Tra 37500 e 41250
20	Tra 41250 e 45000
25	Tra 45000 e 48750
30	Tra 48750 e 52500
35	Tra 52500 e 56250
40	Tra 56250 e 60000
45	Tra 60000 e 63750
50	Tra 63750 e 67500
55	Tra 67500 e 71250

60	Tra 71250 e 75000
65	Tra 75000 e 78750
70	Tra 78750 e 82500
75	Tra 82500 e 86250
80	Tra 86250 e 90000
85	Tra 90000 e 93750
90	Tra 93750 e 97500
95	Tra 97500 e 105000
100	Maggiore di 105000

Nel caso in cui vi siano più registi, ai fini della determinazione della fascia si fa la media aritmetica tra le visualizzazioni e i pesi dei premi vinti da ogni regista.

#### 4.2.2 Fascia di visualizzazioni del film

<b>Fascia</b>	<b>Numero di visualizzazioni</b>
0	Minore di 20000
5	Tra 20000 e 22500
10	Tra 22500 e 25000
15	Tra 25000 e 27500
20	Tra 27500 e 30000
25	Tra 30000 e 32500
30	Tra 32500 e 35000
35	Tra 35000 e 37500
40	Tra 37500 e 40000
45	Tra 40000 e 42500
50	Tra 42500 e 45000
55	Tra 45000 e 47500

60	Tra 47500 e 50000
65	Tra 50000 e 52500
70	Tra 52500 e 55000
75	Tra 55000 e 57500
80	Tra 57500 e 60000
85	Tra 60000 e 62500
90	Tra 62500 e 65000
95	Tra 65000 e 70000
100	Maggiore di 70000

#### 4.2.3 Celebrità degli attori

È data dalla media aritmetica tra la fascia di premialità e la fascia di visualizzazioni dell'attore, calcolate come riportato sotto:

<b>Fascia</b>	<b>Somma dei pesi dei premi ricevuti dall'attore</b>
0	Minore di 5
10	Tra 5 e 10
20	Tra 10 e 15
30	Tra 15 e 20
40	Tra 20 e 25
50	Tra 25 e 30
60	Tra 30 e 35
70	Tra 35 e 40
80	Tra 40 e 45
90	tra 45 e 50
100	Maggiore di 50

<b>Fascia</b>	<b>Numero di visualizzazioni di tutti i film in cui recita l'attore</b>
0	Minore di 40mila
5	tra 40mila e 45mila
10	tra 45mila e 50mila
15	tra 50mila e 55mila
20	Tra 55mila e 60mila
25	Tra 60mila e 65mila
30	Tra 65mila e 70mila
35	Tra 70mila e 75mila
40	Tra 75mila e 80mila
45	Tra 80mila e 85mila
50	Tra 85mila e 90mila
55	Tra 90mila e 95mila
60	Tra 95mila e 100mila
65	Tra 100mila e 105mila
70	Tra 105mila e 110mila
75	Tra 110mila e 115mila
80	Tra 115mila e 120mila
85	Tra 120mila e 125mila
90	Tra 125mila e 130mila
95	Tra 130mila e 140mila
100	Maggiore di 140mila

Nel caso in cui vi siano più attori, ai fini della determinazione della fascia si fa la media aritmetica tra le visualizzazioni e i pesi dei premi vinti da ogni attore. Rammentiamo che le comparse non figurano tra gli attori.

#### 4.2.4 Fascia di premialità del film

Fascia	Somma dei pesi dei premi
0	Minore di 30
10	Tra 30 e 40
20	Tra 40 e 50
30	Tra 50 e 60
40	Tra 60 e 70
50	Tra 70 e 80
60	Tra 80 e 90
70	Tra 90 e 100
80	Tra 100 e 110
90	Tra 110 e 120
100	Maggiore di 120

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Artista	E	2960
Connessione	E	47.45 milioni
Contenuto	E	28mila
Critico	E	30
Direzione	R	600
Erogazione	E	94.9 milioni
Fattore	E	7
Film	E	500
Fruizione	R	94.9 milioni

Importanza	R	600mila
Instaurazione	R	47.45 milioni
Interpretazione	R	5600
Premiazione Attore	R	1120
Premiazione Film	R	750
Premiazione Regista	R	320
Premio	E	2mila
Rappresentazione	R	28mila
Recensione Critico	R	2mila
Recensione Utente	R	300mila
Utente	E	100mila

### Tavola degli accessi

Questa è la prima tavola degli accessi *condizionale* tra quelle che presentiamo. Infatti, dipendentemente dai fattori che l'utente ha indicato in fase di registrazione, è possibile che non sia necessario accedere ad alcune entità o relationship. Abbiamo seguito questo approccio anche in fase di implementazione, accedendo ai dati solo se i fattori dell'utente lo richiedevano, e quindi efficientando le operazioni. Tuttavia, in questa fase preferiamo presentare lo scenario relativo al caso peggiore. Nel caso in esame, il caso peggiore è quello in cui l'utente ha attribuito un valore diverso da 0 a tutti e 7 i fattori.

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Importanza	R	600mila	R	Leggiamo i fattori selezionati dall'utente e il relativo valore.
2	Direzione	R	600	R	Leggiamo i film diretti dai registi del film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
3	Premiazione Regista	R	320	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .

4	Film	E	1	R	Leggiamo la data di uscita del film, per il fattore <i>I film che hanno fatto la storia</i> .
5	Recensione Critico	R	2mila	R	Leggiamo il voto delle recensioni della critica relative al film. Rileva per il fattore <i>Esaltati dalla critica</i> .
6	Recensione Utente	R	600mila	R	Leggiamo il voto delle recensioni degli utenti relative al film. Rileva per il fattore <i>I più amati</i> .
7	Interpretazione	R	5600	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
8	Premiazione Attore	R	1120	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
9	Premiazione Film	R	750	R	Per ogni film, leggiamo quali premi ha ricevuto. Rileva per il fattore <i>Premiati</i> .

10	Premio	E	6.86	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi. Rileva per i fattori <i>Le star</i> , <i>Film d'autore</i> e <i>Premiati</i> .
11	Rappresentazione	R	28mila	R	Leggiamo i contenuti che rappresentano il film di nostro interesse e gli altri 11 film in cui hanno recitato gli attori. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
12	Accesso	R	94.9 milioni	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
13	Fruizione	R	94.9 milioni	R	Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
14	Instaurazione	R	47.45 milioni	R	Per ogni connessione, leggiamo l'utente relativo. Questa azione è fondamentale per ricavare il numero di visualizzazioni. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .



### 4.3 Raccomandazione di contenuti

**Input:** Un utente, rappresentato tramite il suo Id, e un numero naturale  $n$ .

**Output:** Gli identificativi di n film. L'operazione restituisce 0 se nessun film è target per l'utente.

**Frequenza stimata:** 160mila volte al giorno.

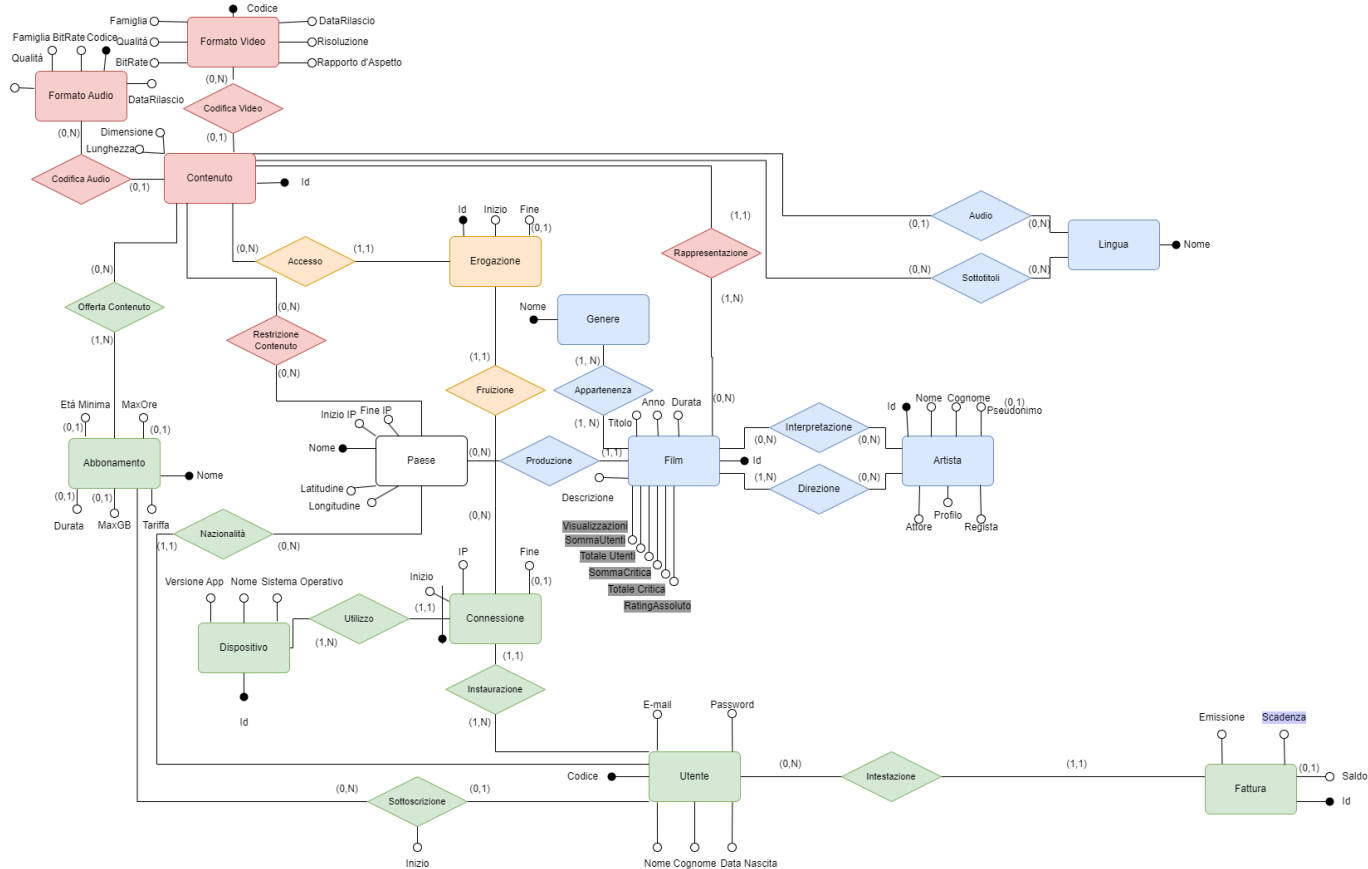


Figura 3: Diagramma ER coinvolto dall'operazione 3. Mancano i formati.

Questa funzione, esplicitamente richiesta nelle *specifiche*, suggerisce all'utente dei film da vedere. A differenza delle funzioni di rating, la raccomandazione dei contenuti si basa unicamente sui film già visualizzati dall'utente. Vengono individuati i seguenti parametri:

Parametro	Peso
Genere	5
Attore	2
Regista	4
Paese di produzione	1

Considerate le visualizzazioni dei film da parte dell'utente, conteggiate come segue, per ogni parametro si assegna a ogni valore un **coefficiente di storico**, calcolato come la percentuale delle visualizzazioni in cui il parametro assume tale valore. Ad esempio, se il 10% delle visualizzazioni di un utente concerne film gialli, a ogni film giallo sarà assegnato un **coefficiente di storico** di 10 per il parametro "genere". In caso di più attori o più registi, viene fatta la media tra i **coefficienti di storico** di ogni attore o ogni regista che ha lavorato al film. Rammentiamo in questa sede che Filmsphere non memorizza le comparse tra gli attori.

Vengono identificati come film target quelli che rispettano le seguenti condizioni:

- Non essere mai stati visualizzati dall'utente;
- Essere fruibili, tramite audio o sottotitoli, in almeno una lingua impiegata dall'utente in almeno il 5% delle sue visualizzazioni;
- Essere fruibili in almeno un formato audio e un formato video impiegati dall'utente in almeno il 5% delle sue visualizzazioni;
- Non essere soggetti a restrizioni nel Paese di residenza dell'utente<sup>11</sup>;
- Essere compresi nel piano di abbonamento dell'utente. L'operazione restituisce una lista vuota se l'abbonamento dell'utente è scaduto o l'utente non è in regola con i pagamenti<sup>12</sup>.

Viene poi fatta, per ogni film target, la media pesata dei **coefficienti di storico** relativi a ogni parametro, utilizzando i pesi riportati in tabella. I film per i quali questa media è più alta vengono restituiti dalla procedura.

Questa operazione è una di quelle che, come anticipato, derogano alla definizione della NOTA 8 in favore di una maggiore efficienza. Nello specifico, i **coefficienti di storico** Genere e Paese sono stati determinati come rapporto tra le erogazioni relative all'utente di contenuti rappresentanti film afferenti al genere o al Paese, e le erogazioni totali relative all'utente; per attori e registi, invece, si è considerato il numero di contenuti diretti o interpretati dall'artista erogati all'utente rispetto al totale.

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Abbonamento	E	5
Accesso	R	94.9 milioni
Appartenenza	R	750
Artista	E	2960
Audio	R	20mila
Codifica Audio	R	20mila
Codifica Video	R	8mila
Connessione	E	47.45 milioni
Contenuto	E	28mila

<sup>11</sup>Si faccia riferimento all'attributo *Nazionalità*.

<sup>12</sup>Consideriamo un utente in regola con i pagamenti se paga le fatture entro un mese dalla loro emissione.

Direzione	R	600
Erogazione	E	94.9 milioni
Fattura	E	780mila
Film	E	500
Formato Audio	E	30
Formato Video	E	60
Fruizione	R	94.9 milioni
Genere	E	10
Instaurazione	R	47.45 milioni
Interpretazione	R	5600
Intestazione	R	780mila
Lingua	E	8
Nazionalità	R	100mila
Offerta Contenuto	R	110mila
Paese	E	40
Produzione	R	500
Rappresentazione	R	28mila
Restrizione Contenuto	R	184800
Sottoscrizione	R	89500
Sottotitoli	R	40mila
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
----	----------	-----------	---------	------	------

1	Instaurazione	R	47.45 milioni	R	Leggiamo le connessioni instaurate dall'utente. Dalla TAVOLA DEI VOLUMI ricaviamo che le connessioni finora instaurate sono 47.45 milioni.
2	Fruizione	R	94.9 milioni	R	Leggiamo le erogazioni relative all'utente. Dalla TAVOLA DEI VOLUMI ricaviamo che le erogazioni finora instaurate sono 94.9 milioni.
3	Accesso	R	94.9 milioni	R	Leggiamo i contenuti finora visualizzati dall'utente.
4	Rappresentazione	R	28mila	R	Leggiamo i film visualizzati dall'utente.
5	Appartenenza	R	750	R	Leggiamo i generi di tutti i film.
6	Interpretazione	R	5600	R	Leggiamo i nomi di tutti gli attori dei vari film.
7	Direzione	R	600	R	Leggiamo i nomi di tutti i registi dei vari film.
8	Produzione	R	500	R	Leggiamo i Paesi dei vari film.
9	Audio	R	20mila	R	Leggiamo le lingue in cui ogni contenuto è doppiato.
10	Sottotitoli	R	40mila	R	Leggiamo le lingue in cui ogni contenuto è sottotitolato.
11	Codifica Audio/Video	R	28mila	R	Leggiamo i formati in cui ogni contenuto audio è codificato.

12	Nazionalità	R	100mila	R	Leggiamo la nazionalità dell'utente. Questa lettura è fondamentale per implementare il vincolo che il film suggerito non sia soggetto a restrizioni nel Paese dell'utente.
13	Restrizione Contenuto	R	184 800	R	-
14	Sottoscrizione	R	89.5mila	R	Leggiamo l'abbonamento sottoscritto dall'utente, oltre a giorno, mese e anno della sottoscrizione.
15	Abbonamento	E	1	R	Leggiamo la durata dell'abbonamento.
16	Offerta Contenuto	R	110mila	R	-
17	Intestazione	R	780mila	R	Questo accesso, con il successivo, è necessario per verificare che l'utente sia in regola con i pagamenti.
18	Fattura	E	7.8	R	A ogni utente attivo viene intestata una fattura ogni mese. Essendo Film-Sphere attivo da esattamente un anno, ed essendo stati gli utenti attivi in media 65mila a ogni utente sono intestate $12 \cdot 0.65$ fatture.

#### 4.4 Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione

**Input:** Un timestamp di inizio connessione, un codice dispositivo e un contenuto, identificato tramite il suo Id.

**Output:** Il server a cui è più conveniente associare la connessione per l'erogazione del contenuto richiesto. Il server è identificato tramite il suo Id.

**Frequenza stimata:** 336mila volte al giorno.

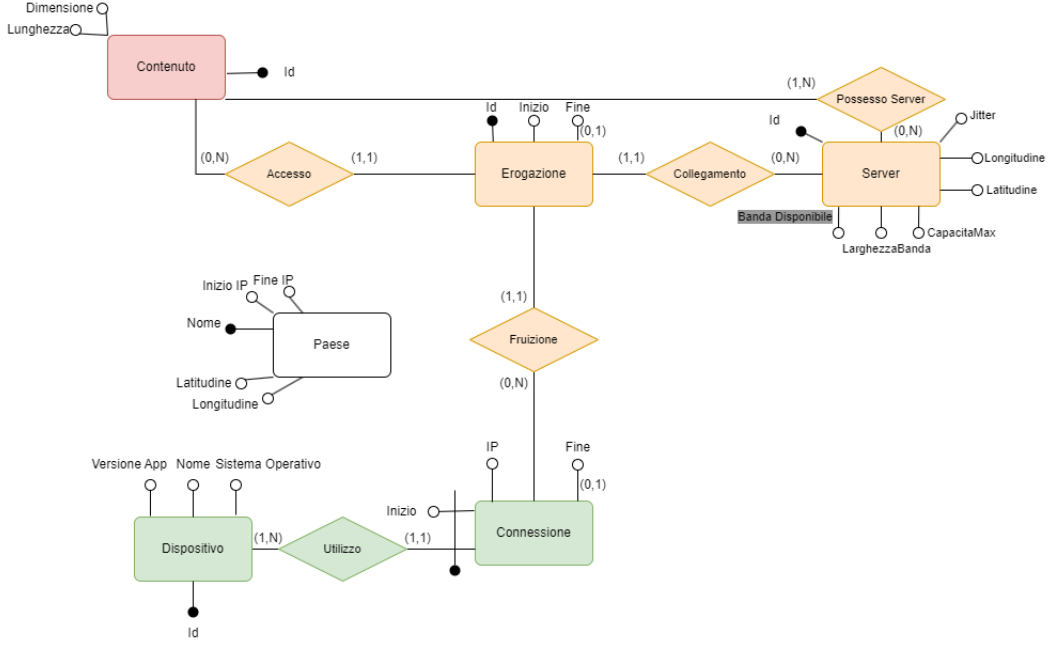


Figura 4: Diagramma ER coinvolto dall'operazione 4.

Quest'operazione, svolta ogni qualvolta un utente richieda la visualizzazione di un contenuto, è vitale per il funzionamento del portale.

Premettiamo alcune note tecniche alla trasmissione di contenuti in streaming, che ci hanno guidato nella formulazione dell'operazione. In primo luogo, il Bit Error Rate è trascurabile: l'errore nella trasmissione è minimo in rapporto alla mole dei dati trasmessi. Possiamo pertanto ignorare, allontanando ben poco il nostro modello dalla realtà, il rapporto segnale-rumore nel teorema fondamentale della teoria dell'informazione, o secondo teorema di Shannon. In secondo luogo, essendo lo streaming in massima parte una trasmissione unidirezionale, la distanza tra l'utente e il server gioca nella comunicazione un ruolo secondario. Infatti, nel protocollo UDP - tipico dei servizi di streaming - la latenza dovuta alla distanza si riflette in un ritardo nell'arrivo del primo pacchetto, ma non aumenta l'intervallo temporale tra due pacchetti contigui, non pregiudicando pertanto la fruizione del servizio. Il reale problema dei servizi di streaming è piuttosto il jitter, ovvero la variabilità nella latenza, di cui abbiamo tenuto conto nell'operazione.

L'operazione calcola, per ogni server, un coefficiente  $\chi$  definito come segue:

$$\chi = \frac{B_{tot} - \sum_{i=1}^n \dim(c_i)/dur(c_i)}{\pi_{Paese} - \pi_{Server}} * \frac{1}{j}$$

dove  $B_{tot}$  è la banda totale del server,  $n$  è il numero di contenuti al momento in erogazione da parte del server,  $\dim(c_i)$  e  $dur(c_i)$  restituiscono, rispettivamente, dimensione (in byte), e durata (in secondi) di un contenuto  $i$ ,  $\pi$  indica la posizione, in coordinate, del baricentro del Paese<sup>13</sup> o del server<sup>14</sup> e  $j$  indica il coefficiente di jitter, a cui si è già accennato sopra. L'utente viene instradato verso il server in cui il coefficiente  $\chi$  è maggiore. L'Id del server viene stampato a video, e viene creata una nuova istanza di erogazione relativa

<sup>13</sup>L'errore di distanza causato dalla non coincidenza dell'utente con le coordinate della nazione è trascurabile per quanto detto in incipit sulla distanza utente-server nel protocollo UDP. Ricordiamo che la localizzazione delle connessioni viene fatta tramite gli intervalli di indirizzi IP assegnati a ogni Paese.

<sup>14</sup>A livello implementativo, le distanze tra due punti sono calcolate con il teorema di Pitagora, considerando le differenze di latitudine e longitudine. Questo metodo non pretende di restituire una distanza reale, dal momento che non tiene conto della geodetica, ma un valore che vi è direttamente proporzionale, e che pertanto è funzionale allo scopo.

alla connessione, al server e al contenuto determinati, con il current timestamp come inizio.

Quest'operazione è computazionalmente molto pesante: abbiamo quindi supposto che, per efficientare Film-Sphere, essa venga invocata solo con parametri di input consistenti, il cui controllo è affidato a operazioni preliminari.

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Collegamento	R	94.9 milioni
Connessione	E	47.45 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Fruizione	R	94.9 milioni
Paese	E	40
Possesso Server	R	100mila
Server	E	10

#### Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Connessione	E	1	R	Leggiamo l'indirizzo IP della connessione data in input.
2	Possesso Server	R	100mila	R	Leggiamo i server che possiedono il contenuto.
3	Server	E	3.57	R	Leggiamo la banda, la posizione e il jitter di ogni server target. Nella TAVOLA DEI VOLUMI abbiamo supposto che ogni server possieda 10mila contenuti in media, ovvero 35.7% del totale. I server sono 10, e ne segue il numero di accessi riportato.

4	Collegamento	R	94.9 milioni	R	Tramite questa relationship si riesce a risalire allo storico delle erogazioni associate ai server target.
5	Erogazione	E	94.9 milioni	R	Leggiamo le erogazioni in corso relative ai server target.
6	Accesso	R	94.9 milioni	R	Per ogni erogazione attiva bisogna stabilire quale contenuto essa stia fornendo, in modo tale da poterne successivamente stabilire il rapporto dimensione/durata.
7	Contenuto	E	4 500	R	Per stabilire il rapporto dimensione/durata del contenuto, è necessario accedere alla rispettiva entità. Stimando che su 10 erogazioni due abbiano stesso contenuto, otteniamo circa 4500 contenuti diversi erogati dalle 5 mila erogazioni target.
8	Paese	E	40	R	Per stabilire la posizione del client è necessario accedere all' entità PAESE e inserire l'IP nel range del paese da cui avviene la connessione.
9	Erogazione	E	1	W	Creiamo una nuova istanza di erogazione, che sia relativa al server, all'utente e alla connessione in oggetto, abbia per inizio il current timestamp e per fine un valore nullo.
10	Accesso	R	1	W	Mettiamo in relazione la nuova erogazione con il contenuto corrispondente.
11	Collegamento	R	1	W	Mettiamo in relazione la nuova erogazione con il server corrispondente.





Categoria	Peso	Note
Generazione	4	Utenti appartenenti alla stessa fascia d'età dell'utente obiettivo. Si veda oltre per un prospetto delle fasce d'età.
Paese	4	Connazionali dell'utente obiettivo.
Preferenze	2	Utenti i cui fattori di preferenza, indicati in fase di registrazione sono affini a quelli dell'utente obiettivo. I fattori di preferenza si dicono affini se non esiste un fattore la cui differenza di valore tra gli utenti superi 20.

Fascia	Anni
Teen	13 - 18
Young	19 - 35
Middle-Aged	35 - 65
Senior	65 +

Definiamo *Movie-Index*  $\mu$  di un contenuto  $\gamma$  la media, pesata sui valori riportati in tabella, delle percentuali di visualizzazioni del film rappresentato da  $\gamma$  da parte degli utenti di ciascuna categoria negli ultimi 14 giorni. Tali percentuali sono calcolate come rapporto tra il numero di contenuti erogati agli utenti negli ultimi 14 giorni che rappresentano il film e il totale dei contenuti erogati agli utenti negli ultimi 14 giorni.

Definiamo poi *Codec-Index*  $\chi$  di un contenuto  $\gamma$  la percentuale di erogazioni, relative all'utente agli ultimi 14 giorni, di contenuti codificati in un formato appartenente alla stessa famiglia di quello in cui è codificato  $\gamma$ .

Definiamo inoltre *Language-Index*  $\lambda$  di un contenuto  $\gamma$  la percentuale di visualizzazioni, relative all'utente e agli ultimi 14 giorni, di contenuti relativi alla stessa lingua di  $\gamma$ . Il *Language-Index* interessa i soli contenuti audio; pertanto viene posto a uno per tutti i contenuti video.

Definiamo infine *Memory-Index*  $\nu$  di un contenuto  $\gamma$  come segue:

$$\nu = \begin{cases} e^{vis(\gamma)-3}, & vis(\gamma) < 3 \\ 0.8^{vis(\gamma)-3}, & vis(\gamma) \geq 3 \end{cases}$$

dove  $vis(\gamma)$  è il numero delle visualizzazioni di  $\gamma$  da parte dell'utente dalla creazione dell'account a oggi. La funzione che utilizziamo è stata da noi selezionata in quanto presenta le seguenti proprietà:

- Cresce esponenzialmente per valori minori di 3: riteniamo infatti che prima di questo valore ogni visualizzazione aumenti significativamente la probabilità di una visualizzazione successiva: l'utente scopre il contenuto e ha piacere di rivederlo;

- Decresce esponenzialmente, ma più lentamente, per valori maggiori di 3: riteniamo infatti che dopo questo valore ogni visualizzazione diminuisca la probabilità di una visualizzazione successiva: l'utente ormai conosce bene il contenuto e ne è stanco.

I parametri della funzione ( $e$  e  $0.8$ ) sono stati scelti arbitrariamente per fornire risultati significativi compatibilmente con il popolamento del database. Si veda la SEZIONE 9 — POPOLAMENTO DEL DATABASE.

A ogni contenuto viene infine associato un indice cumulativo  $\pi = \mu * \chi * \lambda * \nu$ : gli  $n$  contenuti con  $\pi$  maggiore sono restituiti dalla procedura.



Figura 6: Grafico della funzione  $\nu$

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Abbonamento	E	5
Accesso	R	94.9 milioni
Audio	R	20mila
Codifica Audio	R	20mila
Codifica Video	R	8mila
Connessione	E	47.45 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Fattore	E	8
Formato Audio	E	30
Formato Video	E	60
Fruizione	R	94.9 milioni
Importanza	R	600mila
Instaurazione	R	47.45 milioni
Lingua	E	8

Nazionalità	R	100mila
Offerta Contenuto	R	110mila
Paese	E	40
Restrizione Contenuto	R	184800
Sottoscrizione	R	89500
Sottotitoli	R	40mila
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Utente	E	100mila	R	Leggiamo gli utenti appartenenti alla stessa fascia d'età dell'utente obiettivo.
2	Nazionalità	R	100mila	R	Leggiamo gli utenti con nazionalità dell'utente obiettivo. Ipotizziamo in questa sede che gli utenti siano equidistribuiti tra i Paesi.
3	Importanza	R	600mila	R	Leggiamo gli utenti con fattori di preferenza affini a quelli dell'utente obiettivo. Ipotizziamo che, in media, un utente abbia affinità con il 5% della clientela di FilmSphere.
4	Instaurazione	R	47.45 milioni	R	Leggiamo le connessioni relative agli utenti di nostro interesse.
5	Fruizione	R	94.9 milioni	R	Leggiamo le erogazioni relative agli utenti di nostro interesse.

6	Erogazione	E	1575011.5	R	Selezioniamo, per tutti gli utenti tranne quello obiettivo, le sole erogazioni relative agli ultimi 14 giorni. Stimiamo la percentuale di sovrapposizione, ovvero la percentuale di utenti che sono affini all'utente obiettivo in più di una categoria, al 25%. Ipotizziamo che vi siano 25mila utenti della stessa fascia di età dell'utente obiettivo, 2500 utenti suoi connazionali e 30mila utenti con fattori di preferenza a lui affini.
7	Accesso	R	94.9 milioni	R	Leggiamo i contenuti fruiti in ciascuna erogazione.
8	Restrizione Contenuto	R	184800	R	-
9	Sottoscrizione	R	89.5mila	R	Leggiamo l'abbonamento sottoscritto dall'utente, oltre a giorno, mese e anno della sottoscrizione.
10	Abbonamento	E	1	R	Leggiamo la durata dell'abbonamento.
11	Offerta Contenuto	R	110mila	R	-
12	Audio	R	20mila	R	Leggiamo le lingue in cui i contenuti visualizzati dall'utente, e i contenuti target, sono doppiati.
13	Sottotitoli	R	40mila	R	Leggiamo le lingue in cui i contenuti visualizzati dall'utente e i contenuti target sono sottotitolati.
14	Codifica Audio	R	20mila	R	Leggiamo i formati in cui ogni contenuto audio è codificato.
15	Codifica Video	R	8mila	R	Leggiamo i formati in cui ogni contenuto video è codificato.

## 4.6a Registrazione di un utente

**Input:** Un nome, un cognome, un indirizzo e-mail, una password cifrata tramite AES-256, il nome di un Paese e una data di nascita.

**Output:** Un booleano che attesta il successo, o l'insuccesso, dell'operazione.

**Frequenza stimata:** 301.4 volte al giorno.

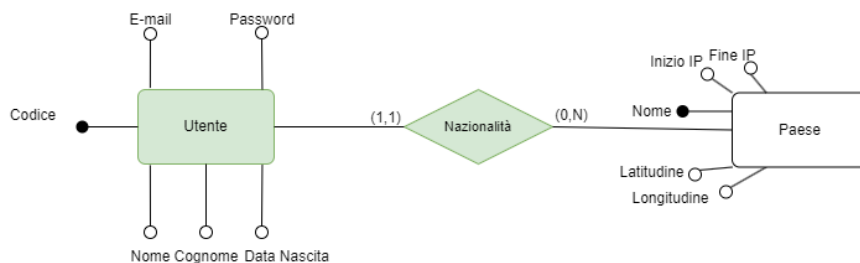


Figura 7: Diagramma ER coinvolto dall'operazione 6a.

La creazione del profilo di un utente si suddivide in due fasi:

- **Registrazione**, ovvero inserimento nel database dei dati dell'utente, comprensivi della nazionalità;
- **Sottoscrizione del servizio**, ovvero scelta del piano di abbonamento e associazione della carta di credito.

La prima fase è trattata in questo paragrafo; la seconda in quello seguente. Il profilo è quindi completato dall'**Espressione delle Preferenze**, non trattata esplicitamente in questa sezione, nella quale l'utente sceglie i fattori di suo interesse e attribuisce loro un valore.

La registrazione degli utenti è un'operazione ad alta frequenza, fondamentale per il ciclo di vita di FilmSphere. Prima della sua esecuzione vengono verificati i seguenti vincoli, i quali, se violati, implicano il fallimento della registrazione e il conseguente ritorno del valore **false**:

- L'indirizzo e-mail deve contenere una chiocciola e un punto;
- Non deve esistere un altro utente con lo stesso indirizzo e-mail (vedasi voce *Utente* nel DIZIONARIO DELLE ENTITÀ);
- Il Paese dell'utente deve essere memorizzato su FilmSphere;<sup>15</sup>
- La data di nascita non deve essere futura.

Il controllo dell'input, sostanzialmente assente nelle operazioni precedenti, è necessario quando i dati sono forniti direttamente dall'utente finale, e non da processi automatizzati o dal personale di FilmSphere. Se l'input è corretto, viene creata una nuova istanza di **UTENTE**. In questo caso supponiamo che lo 0.5% degli utenti si fermi alla prima fase della registrazione, come riportato nella TAVOLA DEI VOLUMI.

### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Nazionalità	R	100mila

<sup>15</sup>Supponiamo che l'utente selezioni il proprio Paese, come per altri campi che ricorreranno di seguito, tramite un menù a tendina. Ciononostante, essendo piuttosto semplice inviare, anche in malafede, richieste al database con dati diversi dalle opzioni presenti nel menù a tendina, abbiamo ritenuto opportuno aggiungere questo controllo ulteriore.

Paese	E	40
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Utente	E	100mila	R	Leggiamo gli utenti con lo stesso indirizzo e-mail dell'utente che andiamo a creare. In caso esistano, l'operazione fallisce.
2	Paese	E	1	R	Verifichiamo che il Paese indicato sia presente su FilmSphere.
3	Utente	E	1	W	Terminate le operazioni di controllo dell'input, creiamo una nuova istanza di utente con i dati forniti.
4	Nazionalità	R	1	W	Creiamo una nuova istanza di nazionalità che metta in relazione l'utente con il suo Paese.

#### 4.6b Sottoscrizione del servizio da parte di un utente

**Input:** Il codice di un utente, il nome di un piano di abbonamento, i dati di una carta di credito (numero, CVV, nome e cognome dell'intestatario, mese e anno di scadenza).

**Output:** Un booleano che attesta il successo, o l'insuccesso, dell'operazione.

**Frequenza stimata:** 298.5 volte al giorno.

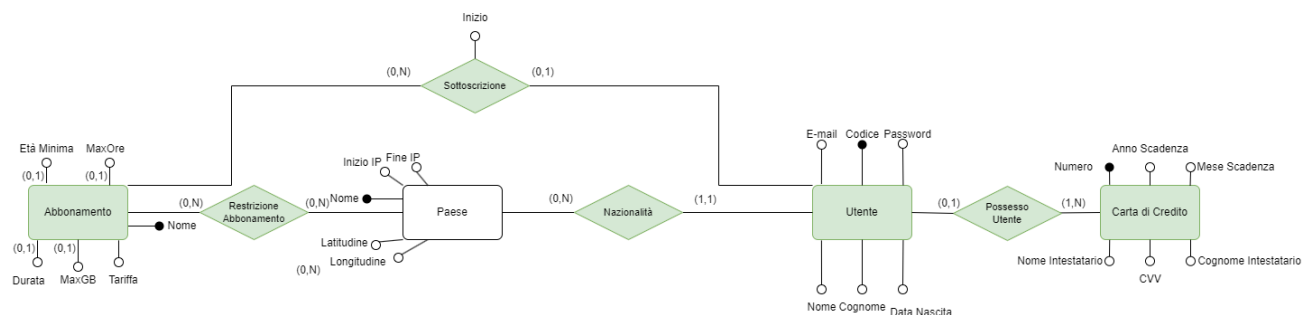


Figura 8: Diagramma ER coinvolto dall'operazione 6b.

Quest'operazione completa la creazione del profilo di un utente. Il codice utente deriva da una procedura automatica e non è pertanto soggetto a controlli, mentre il nome del piano di abbonamento supponiamo sia selezionato dall'utente tramite un menù a tendina, e non tramite un campo testuale. Seguono i vincoli:

- Il numero deve essere composto di esattamente sedici cifre;
- Il CVV deve essere composto esattamente di tre cifre;
- Mese e anno di scadenza devono costituire una data coerente e futura;
- Il piano di abbonamento selezionato deve esistere su FilmSphere;
- L'utente deve rientrare nei limiti di età dell'abbonamento;
- L'abbonamento non deve essere stato interdetto nel Paese dell'utente.

Se l'input è corretto, viene creata, se non già presente, una nuova istanza di CARTA DI CREDITO e associata all'utente. L'utente viene inoltre associato al piano di abbonamento selezionato. Nel caso in cui venga inserito il numero di una carta già presente su FilmSphere con parametri (CVV, dati dell'intestatario, data di scadenza) diversi da quelli forniti, questi non vengono modificati.

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Abbonamento	E	5
Carta di Credito	E	120mila
Nazionalità	R	100mila
Paese	E	40
Possesso Utente	R	99500
Restrizione Abbonamento	R	15
Sottoscrizione	R	89500
Utente	E	100mila

#### Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Abbonamento	E	1	R	Verifichiamo che il piano di abbonamento indicato esista su FilmSphere.
2	Utente	E	1	R	Leggiamo la data di nascita dell'utente allo scopo di calcolarne l'età.



3	Nazionalità	R	100mila	R	Leggiamo il Paese dell'utente.
4	Restrizione Abbonamento	R	15	R	Leggiamo i Paesi in cui è interdetto il piano di abbonamento di nostro interesse.
5	Carta di Credito	E	1	R	Verifichiamo la presenza della carta di credito indicata nel database.
6	Carta di Credito	E	1	W	Nel caso in cui non la trovassimo, creiamo una nuova istanza di carta di credito con i dati forniti in input.
7	Possesso Utente	R	1	W	Associamo la carta di credito all'utente.
8	Sottoscrizione	R	1	W	Associamo l'utente al piano di abbonamento selezionato.

## 4.7 Emissione di una fattura

**Input:** L'identificativo di un utente.

**Output:** Un booleano che attesti il successo, o l'insuccesso dell'operazione.

**Frequenza stimata:** 2630 volte al giorno.

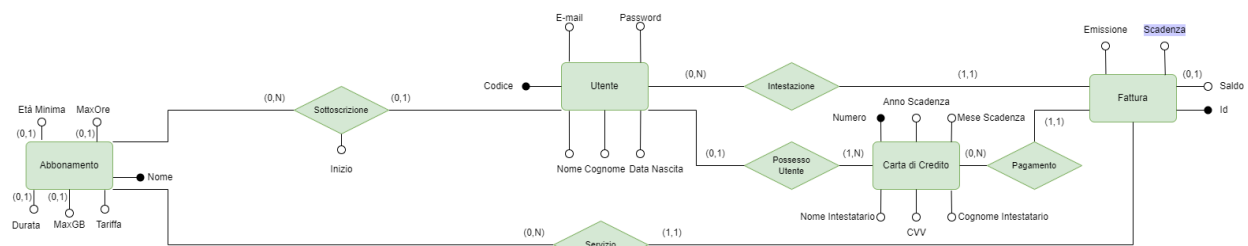


Figura 9: Diagramma ER coinvolto dall'operazione 7.

L'emissione di fattura è gestita tramite un evento: ogni giorno, in un orario in cui la rete è poco carica, viene emessa una fattura per ogni abbonamento che si rinnova quel giorno. Dall'utente si risale al piano di abbonamento e quindi alla carta di credito associata (che deve esserci, vedasi 4.6A REGISTRAZIONE DI UN UTENTE). Viene quindi inserita nel database una nuova fattura, la cui scadenza è fissata a un mese dalla data corrente, e viene posta in relazione con le entità UTENTE, ABBONAMENTO e CARTA DI CREDITO.

L'operazione può fallire in un solo caso, ovvero se la carta di credito dell'utente è scaduta.

**Tavola dei volumi coinvolti**

Concetto	Tipo	Volume
Abbonamento	E	5
Carta di Credito	E	120mila
Fattura	E	780mila
Intestazione	R	780mila
Pagamento	R	780mila
Possesso Utente	R	99500
Servizio	R	780mila
Sottoscrizione	R	89500
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Sottoscrizione	R	89.5mila	R	Leggiamo il piano di abbonamento dell'utente e la sua data di inizio.
2	Abbonamento	E	1	R	Leggiamo la durata dell'abbonamento, in modo da capire se è scaduto o meno. Leggiamo inoltre la tariffa dell'abbonamento.
3	Possesso Utente	R	99.5mila	R	Leggiamo la carta di credito associata all'utente.
4	Carta di Credito	E	1	R	Verifichiamo che la carta di credito non sia scaduta.

5	Fattura	E	1	W	Creiamo una nuova istanza di fattura, con attributo <i>Saldo</i> nullo e con <i>Emissione</i> il current timestamp. Poniamo la <i>Scadenza</i> a un mese dal current timestamp.
6	Pagamento	R	1	W	Associamo la fattura con la carta di credito corrispondente.
7	Intestazione	R	1	W	Associamo la fattura all'utente.
8	Servizio	R	1	W	Associamo la fattura al piano di abbonamento al quale è relativa.

## 4.8 Inserimento di una recensione

**Input:** Un voto espresso in quinti, un testo, l'identificativo di un film, l'identificativo di un utente o di un critico e un booleano che discrimini se l'autore della recensione è un utente o un critico.

**Output:** Un booleano che attesta il successo, o l'insuccesso, dell'operazione.

**Frequenza stimata:** 951.5 volte al giorno.

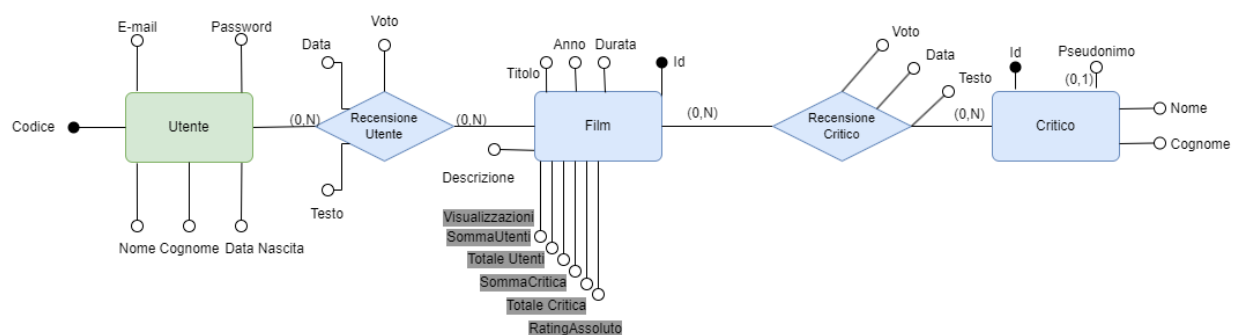


Figura 10: Diagramma ER coinvolto dall'operazione 8.

Le recensioni, come già visto, sono alla base del funzionamento di FilmSphere in quanto partecipano significativamente alle metriche di rating. La data e il booleano di input si suppongono generati automaticamente e pertanto non sono soggetti a controllo. L'operazione fallisce in caso di inserimento di una recensione per un film che l'utente/critico ha già recensito. Questo scenario, che può sfuggire ai controlli lato client ad esempio per motivi di concorrenza, (uno o più dispositivi invocano l'operazione a distanza di pochi millisecondi) deve essere gestito dal database. L'operazione fallisce anche, a causa dei vincoli di cui alla sezione SEZIONE 6.3, nel caso in cui il recensore non sia presente su FilmSphere o il voto non sia compreso tra 1 e 5.

## Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Critico	E	30
Film	E	500
Recensione Critico	R	2mila
Recensione Utente	E	300mila
Utente	R	100mila

#### Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico/Utente	R	1	R	Verifichiamo che non esista una recensione precedente del film da parte del critico/utente.
2	Recensione Critico/Utente	R	1	W	Creiamo una nuova istanza di relationship, che abbia come attributi <i>Voto</i> e <i>Testo</i> i valori forniti, e come attributo <i>Data</i> la current date.

## 4.9 Analytics - Bilanciamento del carico

**Input:** L'identificativo di un server, un intero n.

**Output:** Un numero tra 0 e 100 che indichi il rischio di sovraccarico del server, e una lista, ordinata per priorità decrescente, di n contenuti da spostare in caso di sovraccarico e di server verso cui spostarli.

**Frequenza stimata:** 40 volte l'anno.

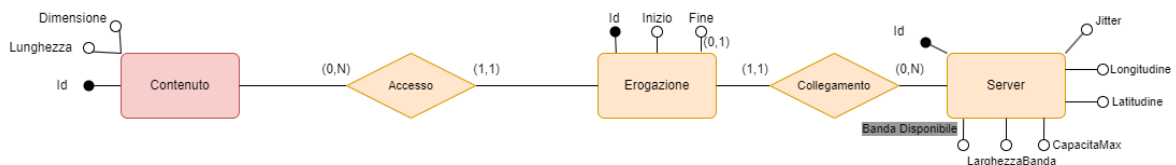


Figura 11: Diagramma ER coinvolto dall'operazione 9.

La trattazione completa dell'operazione è riportata nel PARAGRAFO 8.2.

#### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Collegamento	R	94.9 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Possesso Server	R	100mila
Server	E	10

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Server	E	10	R	Leggiamo gli attributi dei vari server.
2	Erogazione	E	94.9 milioni	R	Leggiamo i timestamp di inizio e di fine di ciascuna erogazione.
3	Collegamento	R	94.9 milioni	R	Poniamo le erogazioni target in relazione con i rispettivi server.
4	Accesso	R	94.9 milioni	R	Leggiamo i contenuti erogati negli ultimi 9 mesi.
5	Contenuto	E	28mila	R	Di ogni contenuto leggiamo la dimensione, allo scopo di determinare la quantità di dati erogata dal server.
6	Possesso Server	R	100mila	R	Questo accesso è necessario per suggerire spostamenti di contenuti.

#### 4.10a Custom Analytics - Fruizione media dei vincoli dell'abbonamento

**Input:** Un mese e un anno.

**Output:** Per ogni fascia d'età, Paese e piano d'abbonamento, la percentuale media di fruizione di contenuti,

in termini di dimensione e di tempo, rispetto ai vincoli dell'abbonamento nel periodo indicato.  
L'operazione:

- Restituisce un insieme vuoto se la data non è consistente con il periodo di attività di FilmSphere. Trattandosi di una procedura eseguita dai dipendenti dell'azienda, preferiamo trascurare, nel solo calcolo degli accessi, questa eventualità;
- Non restituisce risultati relativi ai piani di abbonamento privi di vincoli;
- Non restituisce risultati relativi a combinazioni di piani di abbonamenti, Paesi e fasce d'età a cui non corrispondano erogazioni nel periodo in oggetto.

**Frequenza stimata:** 1 volta al mese.

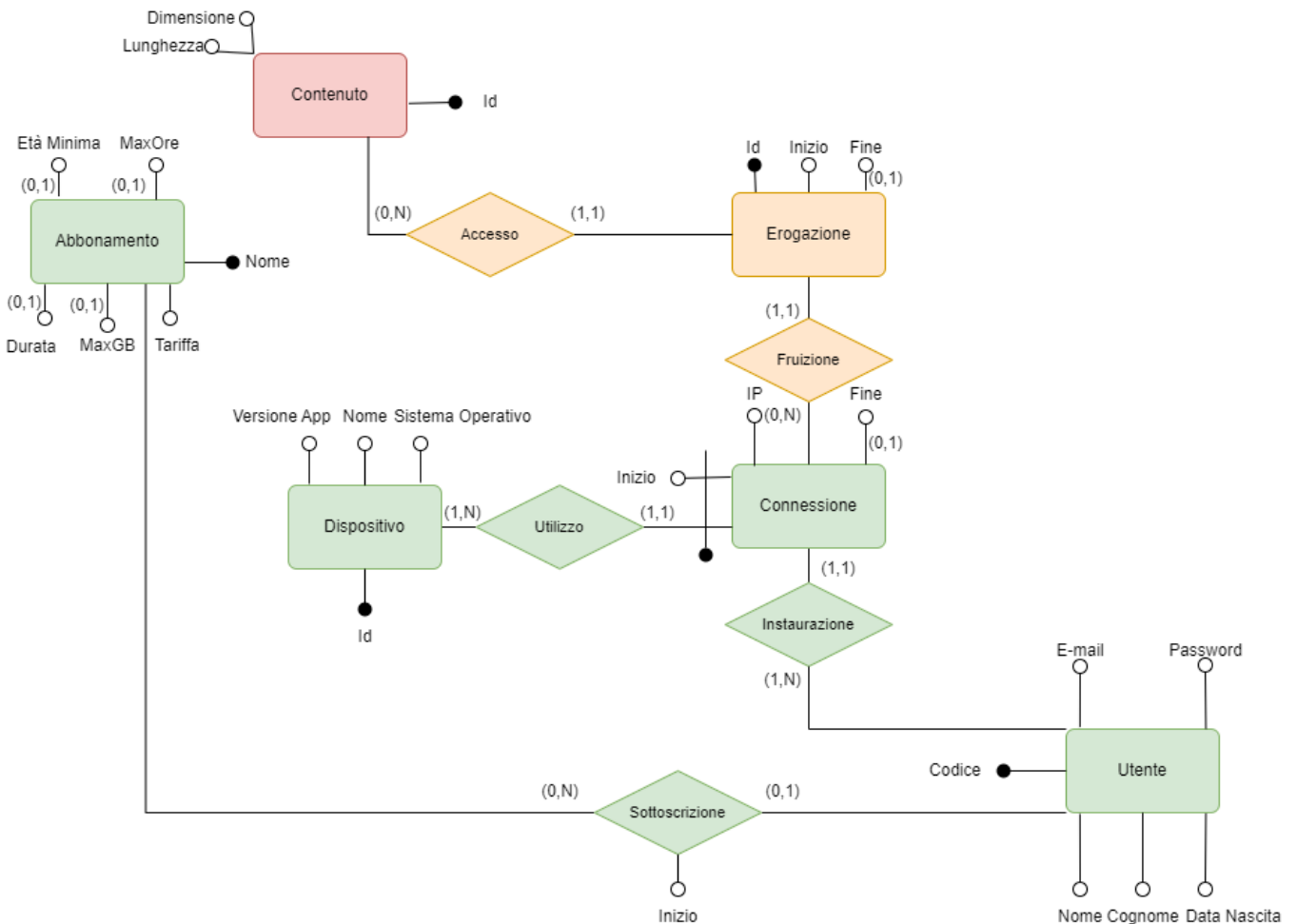


Figura 12: Diagramma ER coinvolto dall'operazione 10a.

La trattazione completa dell'operazione è riportata nel PARAGRAFO 8.3.

Il tempo di visualizzazione di ogni erogazione è stato calcolato come differenza tra i timestamp di fine e inizio, per erogazioni concluse, o come differenza tra il current timestamp e il timestamp di inizio, per erogazioni non concluse. I byte fruiti dall'utente in ogni erogazione sono stati invece determinati come prodotto tra i secondi di erogazione, calcolati come sopra, per la dimensione del contenuto corrispondente in byte divisa per la durata dello stesso in secondi, se la durata dell'erogazione era inferiore alla durata del

contenuto; altrimenti, semplicemente, come la dimensione del contenuto in byte. In questo modo, senza addentrarsi negli algoritmi di caching dei contenuti sul client, si tiene conto solo dei byte effettivamente erogati all'utente, scindendo questo aspetto, ad esempio, dal tempo in cui il film è stato messo in pausa, o dalla velocità di riproduzione.

**Tavola dei volumi coinvolti**

Concetto	Tipo	Volume
Abbonamento	E	5
Accesso	R	47.45 milioni
Connessione	E	94.9 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Fruizione	R	94.9 milioni
Instaurazione	R	47.45 milioni
Nazionalità	R	100mila
Paese	E	40
Sottoscrizione	R	89500
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Erogazione	E	94.9 milioni	R	Leggiamo gli istanti di inizio e di fine di ogni erogazione.
2	Fruizione	R	94.9 milioni	R	Risaliamo alle connessioni relative a ogni erogazione.
3	Instaurazione	R	47.45 milioni	R	Leggiamo gli utenti relativi alle connessioni precedentemente trovate.

4	Sottoscrizione	R	89.5mila	R	Isoliamo gli utenti attivi nel mese e nell'anno indicati, e ne leggiamo i relativi piani di abbonamento.
5	Abbonamento	E	5	R	Leggiamo i limiti, in termini di byte e ore, posti da ogni piano di abbonamento. Stimiamo che solo 3 dei 5 piani pongano limiti di byte e di ore.
6	Accesso	R	94.9 milioni	R	Risaliamo ai contenuti erogati per ogni connessione.
7	Contenuto	E	28mila	R	Leggiamo dimensione e lunghezza di ogni contenuto.
8	Utente	E	65mila	R	Leggiamo la data di nascita degli utenti.

#### 4.10b Custom Analytics - Lingue per tempo di fruizione, come audio e come sottotitoli

**Input:** Nessuno.

**Output:** Per ogni lingua, il numero di ore di riproduzione di contenuti audio e sottotitoli in quella lingua nell'ultimo trimestre, e l'incremento percentuale di tali dati rispetto al trimestre precedente. Le lingue sono identificate tramite il loro nome.

**Frequenza stimata:** 4 volte l'anno.

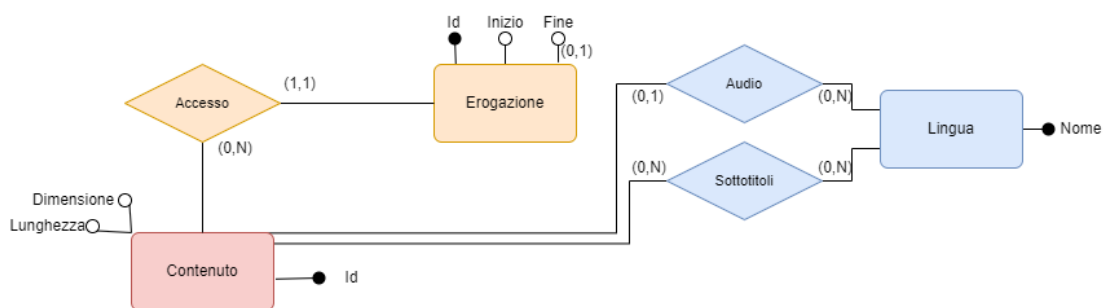


Figura 13: Diagramma ER coinvolto dall'operazione 10b.

La trattazione completa dell'operazione è riportata nel PARAGRAFO 8.4.



**Tavola dei volumi coinvolti**

Concetto	Tipo	Volume
Accesso	R	47.45 milioni
Audio	R	20mila
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Lingua	E	8
Sottotitoli	R	40mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Erogazione	E	94.9 milioni	R	Isoliamo, sulla base degli attributi <i>Inizio</i> e <i>Fine</i> , le erogazioni relative agli ultimi due trimestri.
2	Accesso	R	94.9 milioni	R	Leggiamo tutti i contenuti relativi a ciascuna erogazione.
3	Contenuto	E	28mila	R	Di ogni contenuto leggiamo la durata. Facciamo l'ipotesi, ragionevole, che negli ultimi 6 mesi siano stati erogati almeno una volta tutti i contenuti disponibili sul portale.
4	Sottotitoli	R	40mila	R	Leggiamo le lingue in cui sono sottitolati i vari contenuti.
5	Audio	R	20mila	R	Leggiamo le lingue in cui sono doppiati i vari contenuti.

## 4.11 Analytics - Classifiche

**Input:** Nessuno.

**Output:** Per ogni piano di abbonamento e Paese, la classifica dei film più fruiti dagli utenti nell'ultimo trimestre, con associati i secondi di riproduzione. Per ogni piano di abbonamento, Paese e film la classifica dei formati in cui il film è stato più fruito, con associata la percentuale di fruizione del film. I formati audio e video sono tra loro distinti tramite le lettere A e V.

**Frequenza stimata:** 4 volte l'anno.

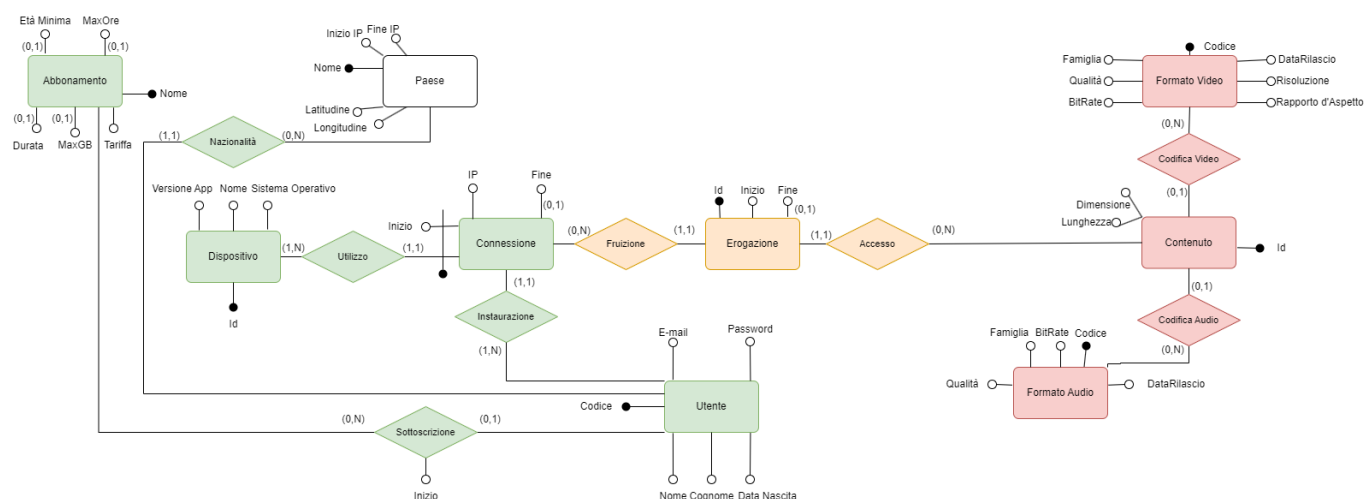


Figura 14: Diagramma ER coinvolto dall'operazione 11.

La trattazione completa dell'operazione è riportata nel PARAGRAFO 8.1.

### Tavola dei volumi coinvolti

Concetto	Tipo	Volume
Abbonamento	E	5
Accesso	R	94.9 milioni
Codifica Audio	R	20mila
Codifica Video	R	8mila
Connessione	E	47.45 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni

Formato Audio	E	30
Formato Video	E	60
Fruizione	R	94.9 milioni
Instaurazione	R	47.45 milioni
Nazionalità	R	100mila
Paese	E	40
Sottoscrizione	R	89500
Utente	E	100mila

**Tavola degli accessi**

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Erogazione	E	94.9 milioni	R	Isoliamo, sulla base degli attributi <i>Inizio</i> e <i>Fine</i> , le erogazioni relative all'ultimo trimestre.
2	Accesso	R	94.9 milioni	R	Leggiamo tutti i contenuti relativi a ciascuna erogazione target.
3	Contenuto	E	28mila	R	Di ogni contenuto leggiamo la lunghezza. Facciamo l'ipotesi, ragionevole, che negli ultimi 3 mesi siano stati erogati, almeno una volta, tutti i contenuti attualmente disponibili sul portale.
4	Codifica Audio	R	20mila	R	Leggiamo i formati in cui sono codificati i contenuti audio.
5	Codifica Video	R	8mila	R	Leggiamo i formati in cui sono codificati i contenuti video.

6	Fruizione	R	94.9 milioni	R	Leggiamo le connessioni relative alle varie erogazioni.
7	Instaurazione	R	47.45 milioni	R	Leggiamo gli utenti relativi. Abbiamo stimato, nella TAVOLA DEI VOLUMI, che per ogni connessione vi siano 2 erogazioni.
8	Sottoscrizione	R	89.5mila	R	Leggiamo gli abbonamenti sottoscritti dagli utenti. Stimiamo un 80% di utenti attivi.
9	Nazionalità	R	100mila	R	Leggiamo i Paesi degli utenti.
10	Rappresentazione	R	28mila	R	Leggiamo i film rappresentati dai vari contenuti.

## 4.12 Fine Erogazione

**Input:** L'Id di un'erogazione e il timestamp corrispondente al momento in cui è terminata.

**Output:** Nessuno.

**Frequenza stimata:** 336 mila volte al giorno.



Figura 15: Diagramma ER coinvolto dall'operazione 12.

Nevralgica nella sua semplicità, questa operazione modifica l'attributo fine dell'erogazione passata come argomento. Sarà essenziale per la formulazione della ridondanza BANDA DISPONIBILE nella prossima sezione.

### Tavola dei volumi coinvolti

Erogazione	E	94.9 milioni
------------	---	--------------

### Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Erogazione	E	1	W	Per ogni erogazione, nell'istante in cui essa termina, è sufficiente un accesso in scrittura alla tabella erogazione, poiché si fa una ricerca sull'attributo chiave, per aggiornare l'attributo Fine.

## 5 — Analisi delle prestazioni delle operazioni

### 5.1 Tavola dei volumi

Entità e relationship di seguito riportate sono relative allo schema ER ristrutturato. Facciamo l'ipotesi, nella stima dei volumi, che FilmSphere sia attivo da esattamente un anno.

Concetto	Tipo	Volume	Nota
Abbonamento	E	5	Come da <i>specifiche</i> . Si rimanda alla SEZIONE 9 per un prospetto completo dei piani d'abbonamento.
Accesso	R	94.9 milioni	Vedasi EROGAZIONE, verso la quale la relationship ha cardinalità (1, 1).
Appartenenza	R	750	Ogni film appartiene in media a 1.5 generi.
Artista	E	2960	Gli artisti possono essere registi, attori o entrambi. Ipotizziamo che l' 8% degli artisti sia contemporaneamente regista e attore. Supponendo che FilmSphere memorizzi i dati di 400 registi e 2800 attori, ne segue che il numero degli artisti è $(400+2800)*100/108 = 2960$ .
Audio	R	20mila	Ogni film è codificato, in media, in 2 contenuti audio per ogni lingua e formato, ed è disponibile mediamente in 5 lingue e 4 formati. I contenuti audio sono pertanto 20000.

Carta di Credito	E	120mila	Considerando gli utenti che non hanno portato a termine la registrazione e quelli che hanno modificato la carta di credito associata al loro profilo, e tenendo conto del fatto che più utenti possono condividere la stessa carta di credito, stimiamo che per ogni utente siano registrate nel sistema 1.2 carte di credito.
Codifica Audio	R	20mila	Poiché ogni contenuto audio è codificato in esattamente un formato audio, il volume della relationship è il medesimo di AUDIO.
Codifica Video	R	8mila	Ogni film è codificato, in media, in 2 contenuti video per ogni formato, ed è disponibile mediamente in 8 formati. I contenuti video sono pertanto 8000.
Collegamento	R	94.9 milioni	Vedasi EROGAZIONE, verso la quale la relationship ha cardinalità (1, 1).
Connessione	E	47.45 milioni	Durante l'anno di attività di Film-Sphere il numero di utenti attivi non è sempre stato costante: alcuni utenti si sono registrati solo negli ultimi mesi; altri hanno disdetto l'abbonamento dopo poco. Stimiamo quindi che in ogni periodo dell'anno siano stati attivi, in media, 65000 utenti. Supponendo che ogni utente attivo si colleghi, in media, al servizio 2 volte al giorno il numero di connessioni è $2 \cdot 365 \cdot 65000 = 47.45$ milioni.
Contenuto	E	28mila	I contenuti possono essere, alternativamente, contenuti video o contenuti audio. I primi sono 8000, i secondi 20000.
Critico	E	30	-
Direzione	R	600	Stimiamo che ogni film sia diretto, in media, da 1.2 registi.

Dispositivo	E	200mila	Tenuto conto del fatto che un dispositivo può essere utilizzato da più utenti, ma anche del fatto che un utente può possedere più dispositivi, stimiamo che siano memorizzati nel database in media 2 dispositivi per utente.
Erogazione	E	94.9 milioni	Stimiamo che ogni giorno ogni utente con un abbonamento attivo visualizzi in media 1 film. Poiché abbiamo supposto che ogni film fosse composto di 4 contenuti, 2 per l'audio e 2 per il video, il numero di erogazioni è $4 \cdot 365 \cdot 65000$ .
Fattore	E	7	Quelli riportati nell'operazione RATING RELATIVO.
Fattura	E	780mila	Per ogni utente che abbia un abbonamento attivo viene emessa una fattura al mese. Abbiamo stimato che il numero di utenti con abbonamento attivo durante l'anno di funzionamento di FilmSphere fosse 65000. Il numero di fatture emesse è pertanto $65000 \cdot 12$ .
Film	E	500	Un valore dimensionato al numero di utenti del portale: Amazon Prime Video, che vanta 200 milioni di utenti (2000 volte quelli di FilmSphere), possiede 26 mila film.
Formato Audio	E	30	Supponiamo che FilmSphere memorizzi 3 famiglie diverse di formati audio, e che a ogni famiglia corrispondano 10 formati.
Formato Video	E	60	Supponiamo che FilmSphere memorizzi 6 famiglie diverse di formati video, e che a ogni famiglia corrispondano 10 formati.
Fruizione	R	94.9 milioni	Vedasi EROGAZIONE, verso la quale la relationship ha cardinalità (1, 1).
Funzionalità Aggiuntiva	E	5	-
Genere	E	10	-

Importanza	R	600mila	Ipotizziamo che ogni utente selezioni, in media, 6 fattori tra i 7 disponibili.
Instaurazione	R	47.45 milioni	Vedasi CONNESSIONE, verso la quale la relationship ha cardinalità (1, 1).
Interpretazione	R	5600	Supponiamo che, in riferimento ai dati di FilmSphere, ogni attore interpreti mediamente 2 film.
Intestazione	R	780mila	Vedasi FATTURA, verso la quale la relationship ha cardinalità (1, 1).
Lingua	E	8	Si è utilizzato come termine di paragone l'azienda leader del settore Netflix, che supporta ad oggi 23 lingue.
Nazionalità	R	100mila	Vedasi UTENTE, verso il quale la relationship ha cardinalità (1, 1).
Offerta Contenuto	R	110mila	Supponiamo che ogni piano di abbonamento dia accesso, in media, a 22mila contenuti. Si rimanda alla SEZIONE 9 per un prospetto completo dei piani d'abbonamento.
Offerta Funzionalità	R	10	Stimiamo che ogni piano di abbonamento offra, in media, 2 funzionalità aggiuntive. Si rimanda alla SEZIONE 9 per un prospetto completo dei piani d'abbonamento.
Paese	E	40	Si è utilizzato come termine di paragone l'azienda leader del settore Netflix, che supporta ad oggi 190 Paesi.
Pagamento	R	780mila	Vedasi FATTURA, verso la quale la relationship ha cardinalità (1, 1).
Possesso Server	R	100mila	Stimiamo che ogni server possieda, in media, 10mila contenuti.



Possesso Utente	R	99500	Ogni utente che abbia completato la seconda fase della registrazione (vedasi 4.6A REGISTRAZIONE DI UN UTENTE) ha associata una carta di credito. Alla scadenza della carta, questa resta associata all'utente finché quest'ultimo non ne inserisce una nuova. Stimiamo che lo 0.5% degli utenti si fermi alla prima fase della registrazione.
Premiazione Attore <sup>16</sup>	R	1120	Considerando che FilmSphere memorizza i dati di tutti gli attori tranne le comparse, stimiamo che un attore abbia in media 0.4 premi.
Premiazione Film	R	750	Stimiamo che un film abbia ricevuto in media 1.5 premi, tenendo conto del fatto che i premi relativi ai film sono sensibilmente di più dei premi relativi ad attori o registi.
Premiazione Regista	R	320	Stimiamo che ogni regista abbia ricevuto in media 0.8 premi.
Premio	E	2mila	La stima tiene conto del fatto che uno stesso premio può essere attribuito a più attori, a più registi, a più film o a una combinazione di questi.
Produzione	R	500	Vedasi FILM, verso il quale la relationship ha cardinalità (1, 1).
Rappresentazione	R	28mila	Vedasi CONTENUTO, verso il quale la relationship ha cardinalità (1, 1).
Recensione Critico	R	2mila	Stimiamo che ogni film del portale sia stato recensito mediamente da 4 critici diversi.

<sup>16</sup>Qui e di seguito è bene considerare che i film presenti su FilmSphere non costituiscono un campione rappresentativo della produzione cinematografica; essi sono film molto richiesti dagli utenti. Conseguenza accidentale di questo fatto è che il numero di premi di cui è insignito un film di FilmSphere è superiore alla media.

Recensione Utente	R	300mila	Stimiamo che ogni utente abbia recensito in media 3 diversi film. La stima tiene conto della notevole dispersione statistica del dato (molti utenti non recensiscono affatto, ma alcuni scrivono decine di recensioni) e del limitato tempo di vita di Film-Sphere, che abbiamo fissato a un anno.
Restrizione Abbonamento	R	15	La stima di questo volume è particolarmente critica. Si rimanda alla SEZIONE 9 per un prospetto dei piani d'abbonamento. Ci aspettiamo, inoltre, un'elevata dispersione statistica: non è improbabile che vi sia un solo piano di abbonamento con restrizioni e che tali restrizioni si applichino a molti Paesi. Possiamo ipotizzare che ogni piano di abbonamento sia interdetto, in media, in 3 Paesi.
Restrizione Contenuto	R	184800	Segmentiamo i contenuti di Film-Sphere in <i>non sensibili</i> , <i>lievemente sensibili</i> , <i>sensibili</i> e <i>altamente sensibili</i> . I contenuti delle varie fasce sono mediamente soggetti a restrizioni, rispettivamente, in 1, 3, 7 e 20 Paesi. Le prime due fasce contengono ognuna il 30% dei contenuti, mentre a ognuna delle ultime due fasce appartiene il 20% dei contenuti. Dalla media pesata segue che ogni contenuto è soggetto a restrizioni, in media, in 6.6 Paesi.
Server	E	10	Un numero proporzionato al numero di Paesi in cui viene erogato Film-Sphere.
Servizio	R	780mila	Vedasi FATTURA, verso la quale la relationship ha cardinalità (1, 1).
Sottoscrizione	R	89500	Stimiamo, come già in POSSESSO UTENTE, che lo 0.5% degli utenti si fermi alla prima fase della registrazione (per la quale si veda 4.6A REGISTRAZIONE DI UN UTENTE), e che il 10% degli utenti abbia un abbonamento scaduto.

Sottotitoli	R	40mila	Stimiamo che ogni contenuto video sia mediamente sottotitolato in 5 lingue.
Utente	E	100mila	La dimensione che stimiamo possa avere un portale di streaming a un anno dal lancio.
Utilizzo	R	47.45 milioni	Vedasi CONNESSIONE, verso la quale la relationship ha cardinalità (1, 1).

## 5.2 Tavola delle operazioni

Le stime che seguono sono un compendio e una giustificazione di quelle riportate nella SEZIONE 4. È fondamentale tenere presente che, stante il continuo ampliamento del bacino d'utenza di FilmSphere, i valori riportati di seguito sono in continua evoluzione. Li abbiamo calcolati in riferimento allo stesso istante in cui abbiamo calcolato i volumi: a un anno dal lancio di FilmSphere. Abbiamo inoltre stimato il numero di utenti attivi in tale istante a 80mila: questo dato non deve essere confuso con il numero di utenti mediamente attivi dal lancio del portale a oggi, che abbiamo fissato a 65mila.

Id	Nome	Frequenza	Nota
1	Rating assoluto, o Valutazione	160mila volte al giorno	L'operazione viene eseguita ogni qualvolta un utente visita la pagina di presentazione di un film, che comprende rating assoluto e relativo. Abbiamo stimato, nella TAVOLA DEI VOLUMI che ogni utente attivo visualizzi in media 1 film al giorno. Supponiamo ora che, per ogni film visualizzato, l'utente visiti la pagina di 2 film. Ne consegue che ogni giorno viene generato il rating assoluto $80000 \cdot 2$ volte.
2	Rating relativo, o Affinità	160mila volte al giorno	Come sopra.

3	Raccomandazione di contenuti	160mila volte al giorno	L'operazione viene eseguita ogni qualvolta un utente si connette a FilmSphere. Il numero di film suggeriti viene calcolato in base al dispositivo, utilizzando una metrica che supponiamo nota. Dall'aver fissato, nella TAVOLA DEI VOLUMI, il numero di connessioni a 2 al giorno per utente attivo, segue la frequenza riportata.
4	Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione	336mila volte al giorno	Abbiamo stimato, nella TAVOLA DEI VOLUMI, che a ogni utente attivo corrispondano 4 erogazioni giornaliere medie. L'operazione viene eseguita a ogni erogazione, per associare l'utente corrispondente a un server. L'operazione viene inoltre eseguita a seguito di un cambio del server CDN durante l'erogazione del contenuto. Stimiamo la probabilità di questo evento al 5%. $80000 \cdot 4 \cdot 105/100 = 336000$ .
5	Caching	80mila volte al giorno	L'operazione viene eseguita una volta al giorno per utente attivo, in un orario in cui il servizio è il più possibile scarico.
6a	Registrazione di un utente	301.4 volte al giorno	In un anno si sono iscritti a FilmSphere 100mila utenti, in media 274 ogni giorno. Stimando al 10% l'aumento del tasso di crescita, rispetto al valor medio, del bacino di utenza del servizio dovuto all'economia di rete, vale $274 \cdot 11/10 = 301.4$ .
6b	Sottoscrizione del servizio da parte di un utente	298.5 volte al giorno	Abbiamo stimato, nella TAVOLA DEI VOLUMI, che lo 0.5% degli utenti completi la prima parte della registrazione ma non la seconda. La stima segue quindi da quella dell'operazione 6.
7	Emissione di una fattura	2630 volte al giorno	Per ogni utente attivo vengono emesse in media 12/365 fatture al giorno. $80000 \cdot 12/365 = 2630$ .

8	Inserimento di una recensione	951.5 volte al giorno	Ad oggi, tra utenti e critici, sono state prodotte 302mila recensioni, 827.4 al giorno. Stimando un tasso di crescita, rispetto al valor medio, del numero delle recensioni del 15% vale $827.4 \cdot 115/100 = 951.5$ , di cui 6.301 recensioni giornaliere dei critici e 945.19 recensioni giornaliere degli utenti.
9	Bilanciamento del carico	40 volte l'anno	L'operazione è eseguita nella programmazione trimestrale degli investimenti aziendali, per ogni server.
10a	Fruizione media dei vincoli dell'abbonamento	1 volta al mese	L'operazione viene eseguita ogni anno in riferimento a tutti i mesi, allo scopo di studiare l'offerta di FilmSphere.
10b	Lingue per tempo di fruizione, come audio e come sottotitoli	4 volte l'anno	L'operazione è eseguita nella programmazione trimestrale degli investimenti aziendali.
11	Classifiche	4 volte l'anno	L'operazione è eseguita nella programmazione trimestrale degli investimenti aziendali.
12	Fine Erogazione	336mila volte al giorno	L'operazione è eseguita per ogni erogazione, e ha pertanto la stessa cardinalità dell'OPERAZIONE 4.

### 5.3 Analisi delle ridondanze

Nella compilazione delle tavole degli accessi di una ridondanza abbiamo considerato operative tutte le ridondanze trattate fino a quel momento. Questo per evidenziare l'efficientamento complessivo di alcune operazioni che sono state semplificate, tramite le ridondanze, per gradi. Gli accessi in scrittura sono stati contati due volte.

#### 5.3.1 Banda Disponibile

La scelta del server più adatto all'erogazione di un certo contenuto ad un determinato utente è un'operazione complessa e ad alta frequenza, il cui efficientamento impatta notevolmente sulla qualità dello streaming di FilmSphere.

Proponiamo di introdurre l'attributo ridondante *Banda Disponibile* nella relazione SERVER, che viene aggiornato in modalità sincrona tramite le operazioni di scelta del server e di fine erogazione.

Questa la frequenza giornaliera delle operazioni che coinvolgono la ridondanza:

Azione	Frequenza Giornaliera
Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione	336 mila
Operazione di fine erogazione	336 mila

Volumi coinvolti (tratti dalla tavola dei volumi alla sezione 5.1)

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Collegamento	R	94.9 milioni
Connessione	E	47.45 milioni
Contenuto	E	28mila
Erogazione	E	94.9 milioni
Fruizione	R	94.9 milioni
Paese	E	40
Possesso Server	R	100mila
Server	E	10

### Tavola degli accessi giornalieri senza ridondanza

Operazione 4 - Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Connessione	E	336mila	R	Leggiamo l'indirizzo IP della connessione data in input.
2	Possesso Server	R	33.6 miliardi	R	Leggiamo i server che possiedono il contenuto.

3	Server	E	1.2 milioni	R	Leggiamo la banda, la posizione e il jitter di ogni server target. Nella TAVOLA DEI VOLUMI abbiamo supposto che ogni server possieda 10mila contenuti in media, ovvero 35.7% del totale. I server sono 10, e ne segue il numero di accessi riportato.
4	Collegamento	R	32 000 miliardi	R	Tramite questa relationship si riesce a risalire allo storico delle erogazioni associate ai server target.
5	Erogazione	E	32 000 miliardi	R	Leggiamo le erogazioni in corso relative ai server target.
6	Accesso	R	32 000 miliardi	R	Per ogni erogazione attiva bisogna stabilire quale contenuto essa stia fornendo, in modo tale da poterne successivamente stabilire il rapporto dimensione/durata. Abbiamo supposto che la media giornaliera di erogazioni sia 336mila e che un contenuto abbia durata media di un'ora. Dividiamo ora per 24 ore del giorno e 10 server, e lo moltiplichiamo per 3.57 server target.
7	Contenuto	E	1.5 miliardi	R	Per stabilire il rapporto dimensione/durata del contenuto, è necessario accedere alla rispettiva entità. Stimando che su 10 erogazioni due abbiano stesso contenuto, otteniamo circa 4500 contenuti diversi erogati dalle 5 mila erogazioni target.

8	Paese	E	13.4 milioni	R	Per stabilire la posizione del client è necessario accedere all' entità PAESE e inserire l'IP nel range del paese da cui avviene la connessione.
9	Erogazione	E	336mila	W	Creiamo una nuova istanza di erogazione, che sia relativa al server, all'utente e alla connessione in oggetto, abbia per inizio il current timestamp e per fine un valore nullo.
10	Accesso	R	336mila	W	Mettiamo in relazione la nuova erogazione con il contenuto corrispondente.
11	Collegamento	R	336mila	W	Mettiamo in relazione la nuova erogazione con il server corrispondente.
12	Fruizione	R	336mila	W	Mettiamo in relazione la nuova erogazione con la connessione corrispondente.
<b>Totale: 96 001 516 608 000 accessi giornalieri</b>					

#### Operazione 12 - Fine Erogazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Erogazione	E	336mila	W	Inseriamo il timestamp di fine erogazione.
<b>Totale: 672mila accessi giornalieri</b>					

**Accessi totali in assenza di ridondanza: 96 001 519 968 000.**

#### Tavola degli accessi giornalieri con la ridondanza

Operazione 4 - Scelta di un server per l'erogazione di un contenuto a un utente, e inizio dell'erogazione



N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Connessione	E	336mila	R	Leggiamo l'indirizzo IP della connessione data in input.
2	Possesso Server	R	33.6 miliardi	R	Leggiamo i server che possiedono il contenuto.
3	Server	E	1.2 milioni	R	Leggiamo la banda, la posizione e il jitter di ogni server target. Nella TAVOLA DEI VOLUMI abbiamo supposto che ogni server possieda 10mila contenuti in media, ovvero 35.7% del totale. I server sono 10, e ne segue il numero di accessi riportato.
4	Collegamento	R	0	R	Questo accesso non è più necessario in quanto le informazioni sulla banda disponibile possono essere ricavate senza accedere alle erogazioni.
5	Erogazione	E	0	R	Le informazioni sulla banda disponibile sono salvate, grazie alla ridondanza, sotto forma di attributo della relazione SERVER.
6	Accesso	R	0	R	Non è più necessario stabilire il rapporto dimensione/durata dei contenuti in erogazione.
7	Contenuto	E	336mila	R	Stabiliamo il rapporto dimensione/durata del contenuto che stiamo per erogare.
8	Paese	E	13.4 milioni	R	Per stabilire la posizione del client è necessario accedere all'entità PAESE e inserire l'IP nel range del paese da cui avviene la connessione.

9	Erogazione	E	336mila	W	Creiamo una nuova istanza di erogazione, che sia relativa al server, all'utente e alla connessione in oggetto, abbia per inizio il current timestamp e per fine un valore nullo.
10	Accesso	R	336mila	W	Mettiamo in relazione la nuova erogazione con il contenuto corrispondente.
11	Collegamento	R	336mila	W	Mettiamo in relazione la nuova erogazione con il server corrispondente.
12	Fruizione	R	336mila	W	Mettiamo in relazione la nuova erogazione con la connessione corrispondente.
13	Server	E	336 mila	W	L'attributo BandaDisponibile necessita di un aggiornamento all'inizio dell'erogazione del contenuto.
<b>Totale: 33 618 632 000 accessi giornalieri</b>					

#### Operazione 12 - Fine Erogazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Accesso	R	32 000 miliardi	R	Risaliamo all'Id del contenuto erogato.
2	Contenuto	E	336 mila	R	Per aggiornare l'attributo ridondante è necessario conoscere il rapporto dimensione/durata del contenuto.
3	Server	E	336 mila	W	L'attributo BandaDisponibile necessita di un aggiornamento alla fine dell'erogazione del contenuto.

4	Erogazione	E	336mila	W	Inseriamo il timestamp di fine erogazione.
<b>Totale: 32 000 001 680 000 accessi giornalieri</b>					

**Accessi totali in presenza di ridondanza: 32 033 620 360 000.**

Si è quindi deciso di mantenere la ridondanza.

### 5.3.2 Rating Assoluto

La valutazione di un film è un'operazione computazionalmente complessa, i cui fattori variano continuamente nel tempo - si pensi alle visualizzazioni o alle recensioni degli utenti - e il cui risultato, tuttavia, subisce minimi cambiamenti quotidiani. Abbiamo quindi pensato di introdurre, nell'entità FILM, l'attributo ridondante *Rating Assoluto*, da aggiornare quotidianamente in modalità deferred. In questo modo accettiamo di restituire all'utente un risultato calcolato su dati vecchi di qualche ora, ma efficientiamo notevolmente l'operazione.

Azione	Frequenza Giornaliera
Visualizzazione della pagina di un film da parte di un utente, e conseguente calcolo del <i>Rating Assoluto</i>	160 mila

Volumi coinvolti (tratti dalla tavola dei volumi alla sezione 5.1)

Concetto	Costrutto	Volume
Film	E	500
Interpretazione	R	5 600
Direzione	R	600
Artista	E	2 960
PremiazioneRegista	R	320
PremiazioneAttore	R	1 120
PremiazioneFilm	R	750
Premio	E	2 000
RecensioneCritico	R	2 000

RecensioneUtente	R	300 mila
Critico	E	30
Utente	E	100 mila
Rappresentazione	R	28 mila
Contenuto	E	28 mila
Accesso	R	94.9 milioni
Erogazione	E	94.9 milioni
Fruizione	R	94.9 milioni
Connessione	E	47.45 milioni
Instaurazione	R	47.45 milioni

### Tavola degli accessi in assenza di ridondanza

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico	R	320 milioni	R	Leggiamo il voto delle recensioni della critica relative al film.
2	Recensione Utente	R	48 miliardi	R	Leggiamo il voto delle recensioni degli utenti relative al film.
3	Premiazione Film	R	120 milioni	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	896 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	96 milioni	R	Leggiamo i film diretti dai registi del film di nostro interesse.

6	Premiazione Attore	R	179.2 milioni	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	51.2 milioni	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.
8	Premio	E	1.1 milioni	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.
9	Rappresentazione	R	4.5 miliardi	R	Leggiamo i contenuti che rappresentano il film di nostro interesse e gli altri 11 film in cui hanno recitato gli attori.
10	Accesso	R	15 000 miliardi	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto.
11	Fruizione	R	15 000 miliardi	R	Per ogni erogazione, leggiamo la connessione destinataria.
12	Instaurazione	R	7 500 miliardi	R	Per ogni connessione, leggiamo l'utente relativo. Questa azione è fondamentale per ricavare il numero di visualizzazioni.
<b>Totale: 37 554 163 500 000 accessi giornalieri</b>					

## Tavola degli accessi in presenza di ridondanza

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Film	E	160 mila	R	È sufficiente accedere alla relazione Film e visualizzare l'attributo RatingAssoluto
<b>Totale: 160 000 accessi giornalieri</b>					

Per l'aggiornamento deferred con frequenza giornaliera si svolge la stessa procedura discussa nella parte senza ridondanza, con un numero di accessi diviso per 160 mila / 500 = 320, dal momento che l'operazione viene eseguita una volta al giorno per ogni film presente su FilmSphere. Sono necessari inoltre 500 accessi giornalieri in scrittura per aggiornare la relazione film. Gli accessi giornalieri in presenza di ridondanza sono quindi **117 356 922 000**.

### 5.3.3 Visualizzazioni

Un fattore importante nelle procedure di Rating è il numero di visualizzazioni, il cui calcolo è computazionalmente complesso. Proponiamo l'introduzione di un attributo *Visualizzazioni* sull'entità FILM che sia aggiornato in modalità asincrona, con cadenza giornaliera. Per la definizione di visualizzazione rimandiamo alla NOTA 8.

Di seguito è riportata la frequenza giornaliera delle attività che coinvolgono tale ridondanza

Azione	Frequenza Giornaliera
Rating Assoluto, o Valutazione	500
Rating Relativo, o Affinità	160 mila

Volumi coinvolti (tratti dalla tavola dei volumi alla sezione 5.1)

Concetto	Tipo	Volume
Accesso	R	94.9 milioni
Artista	E	2960
Collegamento	R	94.9 milioni
Connessione	E	47.45 milioni
Contenuto	E	28mila

Critico	E	30
Direzione	R	600
Erogazione	E	94.9 milioni
Fattore	E	7
Film	E	500
Fruizione	R	94.9 milioni
Importanza	R	600mila
Instaurazione	R	47.45 milioni
Interpretazione	R	5600
Paese	E	40
Possesso Server	R	100mila
Premiazione Attore	R	1120
Premiazione Film	R	750
Premiazione Regista	R	320
Premio	E	2mila
Rappresentazione	R	28mila
Recensione Critico	R	2mila
Recensione Utente	R	300mila
Server	E	10
Utente	E	100mila

#### Tavola degli accessi in assenza di ridondanza

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
----	----------	-----------	---------	------	------

1	Recensione Critico	R	1 milione	R	Leggiamo il voto delle recensioni della critica relative al film.
2	Recensione Utente	R	150 milioni	R	Leggiamo il voto delle recensioni degli utenti relative al film.
3	Premiazione Film	R	375mila	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	2.8 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	300mila	R	Leggiamo i film diretti dai registi del film di nostro interesse.
6	Premiazione Attore	R	560mila	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	160mila	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.
8	Premio	E	3430	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.



9	Rappresentazione	R	14 milioni	R	Leggiamo i contenuti che rappresentano il film di nostro interesse e gli altri 11 film in cui hanno recitato gli attori.
10	Accesso	R	47.5 miliardi	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto.
11	Fruizione	R	47.5 miliardi	R	Per ogni erogazione, leggiamo la connessione destinataria.
12	Instaurazione	R	23.75 miliardi	R	Per ogni connessione, leggiamo l'utente relativo. Questa azione è fondamentale per ricavare il numero di visualizzazioni.
13	Film	E	500	W	Aggiorniamo l'attributo ridondante Rating Assoluto.
<b>Totale: 118 919 199 000 accessi giornalieri</b>					

Operazione 2 - Rating Relativo, o Affinità

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Importanza	R	96 miliardi	R	Leggiamo i fattori selezionati dall'utente e il relativo valore.
2	Direzione	R	96 milioni	R	Leggiamo i film diretti dai registi del film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
3	Premiazione Regista	R	51.2 milioni	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .

4	Film	E	160mila	R	Leggiamo la data di uscita del film, per il fattore <i>I film che hanno fatto la storia</i> .
5	Recensione Critico	R	320 milioni	R	Leggiamo il voto delle recensioni della critica relative al film. Rileva per il fattore <i>Esaltati dalla critica</i> .
6	Recensione Utente	R	96 miliardi	R	Leggiamo il voto delle recensioni degli utenti relative al film. Rileva per il fattore <i>I più amati</i> .
7	Interpretazione	R	896 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
8	Premiazione Attore	R	179.2 milioni	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
9	Premiazione Film	R	120 milioni	R	Per ogni film, leggiamo quali premi ha ricevuto. Rileva per il fattore <i>Premiati</i> .

10	Premio	E	1.1 milioni	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi. Rileva per i fattori <i>Le star</i> , <i>Film d'autore</i> e <i>Premiati</i> .
11	Rappresentazione	R	4.5 miliardi	R	Leggiamo i contenuti che rappresentano il film di nostro interesse e gli altri 11 film in cui hanno recitato gli attori. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
12	Accesso	R	15 000 miliardi	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
13	Fruizione	R	15 000 miliardi	R	Vedi accesso 11 del RATING ASSOLUTO. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
14	Instaurazione	R	7 500 miliardi	R	Per ogni connessione, leggiamo l'utente relativo. Questa azione è fondamentale per ricavare il numero di visualizzazioni. Rileva per i fattori <i>Popolari</i> , <i>Le star</i> , <i>Film d'autore</i> .
<b>Totale: 37 698 163 700 000 accessi giornalieri</b>					

**Accessi totali in assenza di ridondanza: 37 817 082 901 000.**

Per mantenere aggiornata questa ridondanza è necessario accedere in scrittura all'entità FILM e modificare l'attributo in questione all'inizio di ogni erogazione del contenuto che si qualifichi come visualizzazione.

### **Tavola degli accessi in presenza di ridondanza**

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico	R	1 milione	R	Leggiamo il voto delle recensioni della critica relative al film.
2	Recensione Utente	R	150 milioni	R	Leggiamo il voto delle recensioni degli utenti relative al film.
3	Premiazione Film	R	375mila	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	2.8 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	300mila	R	Leggiamo i film diretti dai registi del film di nostro interesse.
6	Premiazione Attore	R	560mila	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	160mila	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.

8	Premio	E	3430	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.
9	Film	E	6300	R	In questa sede siamo interessati al film target, ai film in cui recitano gli attori del film target e ai film diretti dai registi del film target.
10	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
11	Accesso	R	0	R	-
12	Fruizione	R	0	R	-
13	Instaurazione	R	0	R	-
14	Film	E	500	W	Aggiorniamo l'attributo ridondante Rating Assoluto.
<b>Totale: 155 205 730 accessi giornalieri</b>					

Operazione 2 - Rating Relativo, o Affinità

N°	Concetto	Costrutto	Accessi	Tipo	Note
----	----------	-----------	---------	------	------

1	Importanza	R	96 miliardi	R	Leggiamo i fattori selezionati dall'utente e il relativo valore.
2	Direzione	R	96 milioni	R	Leggiamo i film diretti dai registi del film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
3	Premiazione Regista	R	51.2 milioni	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
4	Film	E	2 milioni	R	Leggiamo le visualizzazioni del film target, dei film diretti dai registi del film target e dei film in cui recitano gli attori del film target. Leggiamo anche la data di uscita del film target, per il fattore <i>I film che hanno fatto la storia</i> .
5	Recensione Critico	R	320 milioni	R	Leggiamo il voto delle recensioni della critica relative al film. Rileva per il fattore <i>Esaltati dalla critica</i> .
6	Recensione Utente	R	96 miliardi	R	Leggiamo il voto delle recensioni degli utenti relative al film. Rileva per il fattore <i>I più amati</i> .
7	Interpretazione	R	896 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
8	Premiazione Attore	R	179.2 milioni	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .

9	Premiazione Film	R	120 milioni	R	Per ogni film, leggiamo quali premi ha ricevuto. Rileva per il fattore <i>Premiati</i> .
10	Premio	E	1.1 milioni	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi. Rileva per i fattori <i>Le star</i> , <i>Film d'autore</i> e <i>Premiati</i> .
11	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
12	Accesso	R	0	R	-
13	Fruizione	R	0	R	-
14	Instaurazione	R	0	R	-
<b>Totale: 193 665 500 000 accessi giornalieri</b>					

Aggiorniamo la ridondanza quotidianamente tramite un evento. Durante l'aggiornamento vengono lette per intero tutte le relazioni coinvolte e vengono aggiornate simultaneamente tutte le istanze dell'entità film.

Aggiornamento della ridondanza

N°	Concetto	Costrutto	Accessi	Tipo	Note
----	----------	-----------	---------	------	------

1	Rappresentazione	R	28mila	R	Associamo ogni contenuto all'Id del rispettivo film.
2	Accesso	R	94.9 milioni	R	Per ogni contenuto leggiamo le erogazioni nelle quali è coinvolto.
3	Fruizione	R	94.9 milioni	R	Per ogni erogazione, leggiamo la connessione destinataria.
4	Instaurazione	R	47.45 milioni	R	Per ogni connessione, leggiamo l'utente relativo.
5	Film	E	500	W	Aggiorniamo l'attributo visualizzazioni per ogni film.
<b>Totale: 237 279 000 accessi giornalieri</b>					

**Accessi totali in presenza di ridondanza: 194 057 985 000.**

Si è quindi deciso di mantenere la ridondanza.

#### **5.3.4 Somma Recensioni della Critica, Totale Recensioni della Critica, Somma Recensioni degli Utenti e Totale Recensioni degli Utenti**

La media delle recensioni della critica svolge un ruolo molto importante nelle OPERAZIONI 4.1 e 4.2, che coinvolgono la determinazione, rispettivamente, del rating assoluto e relativo di un film. Queste ultime sono operazioni complesse e ad alta frequenza, la cui ottimizzazione è cruciale per l'efficienza di FilmSphere. Di seguito proponiamo di aggiungere quattro attributi ridondanti all'entità FILM, denominati *Somma Critica*, *Somma Utenti*, *Recensioni Critica* e *Recensioni Utenti* che rappresentino, rispettivamente, la somma dei valori di tutte le recensioni che il film ha ricevuto, da parte della critica e degli utenti, e il numero di tali recensioni. La ridondanza si aggiornerebbe tramite immediate refresh all'atto dell'inserimento di una recensione.

Di seguito è riportata la frequenza giornaliera delle attività che coinvolgono tale ridondanza.

<b>Azione</b>	<b>Frequenza Giornaliera</b>
Rating Assoluto, o Valutazione	160mila
Rating Relativo, o Affinità	160mila
Inserimento di una recensione	951.5



Volumi coinvolti (tratti dalla tavola dei volumi alla sezione 5.1)

<b>Concetto</b>	<b>Tipo</b>	<b>Volume</b>
Accesso	R	94.9 milioni
Artista	E	2960
Connessione	E	47.45 milioni
Contenuto	E	28mila
Critico	E	30
Direzione	R	600
Erogazione	E	94.9 milioni
Fattore	E	7
Film	E	500
Fruizione	R	94.9 milioni
Importanza	R	600mila
Instaurazione	R	47.45 milioni
Interpretazione	R	5600
Premiazione Attore	R	1120
Premiazione Film	R	750
Premiazione Regista	R	320
Premio	E	2mila
Rappresentazione	R	28mila
Recensione Critico	R	2mila
Recensione Utente	R	300mila
Utente	E	100mila

#### **Tavole degli accessi in assenza di ridondanza**

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico	R	1 milione	R	Leggiamo il voto delle recensioni della critica relative al film.
2	Recensione Utente	R	150 milioni	R	Leggiamo il voto delle recensioni degli utenti relative al film.
3	Premiazione Film	R	375mila	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	2.8 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	300mila	R	Leggiamo i film diretti dai registi del film di nostro interesse.
6	Premiazione Attore	R	560mila	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	160mila	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.
8	Premio	E	3430	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.

9	Film	E	6300	R	In questa sede siamo interessati al film target, ai film in cui recitano gli attori del film target e ai film diretti dai registi del film target.
10	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
11	Accesso	R	0	R	-
12	Fruizione	R	0	R	-
13	Instaurazione	R	0	R	-
14	Film	E	500	W	Aggiorniamo l'attributo ridondante Rating Assoluto.
<b>Totale: 155 205 730 accessi giornalieri</b>					

Operazione 2 - Rating Relativo, o Affinità

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Importanza	R	96 miliardi	R	Leggiamo i fattori selezionati dall'utente e il relativo valore.
2	Direzione	R	96 milioni	R	Leggiamo i film diretti dai registi del film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
3	Premiazione Regista	R	51.2 milioni	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .

4	Film	E	2 milioni	R	Leggiamo le visualizzazioni del film target, dei film diretti dai registi del film target e dei film in cui recitano gli attori del film target. Leggiamo anche la data di uscita del film target, per il fattore <i>I film che hanno fatto la storia.</i>
5	Recensione Critico	R	320 milioni	R	Leggiamo il voto delle recensioni della critica relative al film. Rileva per il fattore <i>Esaltati dalla critica.</i>
6	Recensione Utente	R	96 miliardi	R	Leggiamo il voto delle recensioni degli utenti relative al film. Rileva per il fattore <i>I più amati.</i>
7	Interpretazione	R	896 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star.</i>
8	Premiazione Attore	R	179.2 milioni	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star.</i>
9	Premiazione Film	R	120 milioni	R	Per ogni film, leggiamo quali premi ha ricevuto. Rileva per il fattore <i>Premiati.</i>

10	Premio	E	1.1 milioni	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi. Rileva per i fattori <i>Le star</i> , <i>Film d'autore</i> e <i>Premiati</i> .
11	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
12	Accesso	R	0	R	-
13	Fruizione	R	0	R	-
14	Instaurazione	R	0	R	-
<b>Totale: 193 665 500 000 accessi giornalieri</b>					

Operazione 8 - Inserimento di una recensione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico/Utente	R	951.5	R	Verifichiamo che non esista una recensione precedente del film da parte del critico/utente.

2	Recensione Critico/Utente	R	951.5	W	Creiamo una nuova istanza di relationship, che abbia come attributi <i>Voto</i> e <i>Testo</i> i valori forniti, e come attributo <i>Data</i> la current date.
<b>Totale: 2 854.5 accessi giornalieri</b>					

**Accessi totali in assenza di ridondanza: 193 820 701 951.**

#### **Tavole degli accessi in presenza di ridondanza**

Operazione 1 - Rating Assoluto, o Valutazione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico	R	0	R	La media delle recensioni della critica è ricavabile dagli attributi <i>Somma Critica</i> e <i>Totale Critica</i> dell'entità FILM.
2	Recensione Utente	R	0	R	La media delle recensioni degli utenti è ricavabile dagli attributi <i>Somma Utenti</i> e <i>Totale Utenti</i> dell'entità FILM.
3	Premiazione Film	R	375mila	R	Per ogni film, leggiamo quali premi ha ricevuto.
4	Interpretazione	R	2.8 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse.
5	Direzione	R	300mila	R	Leggiamo i film diretti dai registi del film di nostro interesse.
6	Premiazione Attore	R	560mila	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse.
7	Premiazione Regista	R	160mila	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse.

8	Premio	E	3430	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi.
9	Film	E	6300	R	In questa sede siamo interessati al film target, ai film in cui recitano gli attori del film target e ai film diretti dai registi del film target. Leggiamo anche gli attributi <i>Somma Critica</i> , <i>Totale Critica</i> , <i>Somma Utenti</i> e <i>Totale Utenti</i> .
10	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
11	Accesso	R	0	R	-
12	Fruizione	R	0	R	-
13	Instaurazione	R	0	R	-
14	Film	E	500	W	Aggiorniamo l'attributo ridondante Rating Assoluto.
<b>Totale: 4 205 730 accessi giornalieri</b>					

Operazione 2 - Rating Relativo, o Affinità

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Importanza	R	96 miliardi	R	Leggiamo i fattori selezionati dall'utente e il relativo valore.
2	Direzione	R	96 milioni	R	Leggiamo i film diretti dai registi del film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
3	Premiazione Regista	R	51.2 milioni	R	Leggiamo i premi ricevuti dai registi presenti nel film di nostro interesse. Rileva per il fattore <i>Film d'autore</i> .
4	Film	E	2 milioni	R	Leggiamo le visualizzazioni del film target, dei film diretti dai registi del film target e dei film in cui recitano gli attori del film target. Leggiamo inoltre gli attributi <i>Somma Critica</i> , <i>Totale Critica</i> , <i>Somma Utenti</i> e <i>Totale Utenti</i> . Leggiamo infine la data di uscita del film target, per il fattore <i>I film che hanno fatto la storia</i> .
5	Recensione Critico	R	0	R	La media delle recensioni della critica è ricavabile dagli attributi <i>Somma Critica</i> e <i>Totale Critica</i> dell'entità FILM.
6	Recensione Utente	R	0	R	La media delle recensioni degli utenti è ricavabile dagli attributi <i>Somma Utenti</i> e <i>Totale Utenti</i> dell'entità FILM.
7	Interpretazione	R	896 milioni	R	Leggiamo i film in cui recitano gli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .



8	Premiazione Attore	R	179.2 milioni	R	Leggiamo i premi ricevuti dagli attori presenti nel film di nostro interesse. Rileva per il fattore <i>Le star</i> .
9	Premiazione Film	R	120 milioni	R	Per ogni film, leggiamo quali premi ha ricevuto. Rileva per il fattore <i>Premiati</i> .
10	Premio	E	1.1 milioni	R	Leggiamo i pesi dei premi ricevuti da attori, registi e film. Ogni attore riceve in media 0.4 premi; in ogni film recitano mediamente 11 attori. Il numero di premi ricevuti da attori che interpretano il film di nostro interesse è quindi 4.4. Ogni film è diretto in media da 1.2 registi e ogni regista riceve mediamente 0.8 premi. Il numero di premi mediamente attribuiti ai registi del film di nostro interesse è quindi 0.96. Infine stimiamo che ogni film sia insignito di 1.5 premi. Rileva per i fattori <i>Le star</i> , <i>Film d'autore</i> e <i>Premiati</i> .
11	Rappresentazione	R	0	R	Qui e di seguito, non è necessario accedere direttamente alle erogazioni dei contenuti che rappresentano i vari film per determinarne le visualizzazioni.
12	Accesso	R	0	R	-
13	Fruizione	R	0	R	-
14	Instaurazione	R	0	R	-
<b>Totale: 97 345 500 000 accessi giornalieri</b>					

Operazione 8 - Inserimento di una recensione

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Recensione Critico/Utente	R	951.5	R	Verifichiamo che non esista una recensione precedente del film da parte del critico/utente.
2	Recensione Critico/Utente	R	951.5	W	Creiamo una nuova istanza di relationship, che abbia come attributi <i>Voto</i> e <i>Testo</i> i valori forniti, e come attributo <i>Data</i> la current date.
3	Film	E	951.5	W	Aggiorniamo gli attributi <i>Somma Critica</i> e <i>Totale Critica</i> , o <i>Somma Utenti</i> e <i>Totale Utenti</i> .
<b>Totale: 4 757.5 accessi giornalieri</b>					

**Accessi totali in presenza di ridondanza: 97 349 710 503.**

Si è quindi deciso di mantenere la ridondanza.

### 5.3.5 Tabella riepilogativa

Al termine della nostra analisi, lo schema Entity-Relationship risulta modificato come segue:

Attributo ridondante	Entità di riferimento	Refresh
Banda Disponibile	Server	Immediate
Rating Assoluto	Film	Deferred
Visualizzazioni	Film	Deferred
Somma Critica	Film	Immediate
Totale Critica	Film	Immediate
Somma Utenti	Film	Immediate
Totale Utenti	Film	Immediate

## 6 — Progettazione Logica

### 6.1 Descrizione dello schema logico

La traduzione dello schema concettuale in schema logico produce le seguenti tabelle:

Abbonamento(Nome, Tariffa, Durata, MaxOre, EtaMinima, MaxGB)  
Appartenenza(Film, Genere)  
Artista(Id, Nome, Cognome, Pseudonimo, Profilo, Attore, Regista)  
CartaDiCredito(Numero, CVV, NomeIntestatario, CognomeIntestatario, MeseScadenza, AnnoScadenza)  
CodificaVideo(Contenuto, FormatoVideo)  
Connessione(Inizio, Dispositivo, IP, Fine, Utente)  
Contenuto(Id, Dimensione, Lunghezza, Film, LinguaAudio, CodificaAudio)<sup>17</sup>  
Critico(Id, Nome, Cognome, Pseudonimo)  
Direzione(Film, Artista)  
Dispositivo(Id, Nome, VersioneApp, SistemaOperativo)  
Erogazione(Id, Inizio, Fine, Contenuto, Server, InizioConnessione, Dispositivo)  
Fattura(Id, Saldo, Scadenza, Emissione, Utente, CartaDiCredito, Abbonamento)  
Film(Id, Titolo, Descrizione, Anno, Durata, Paese, SommaCritica, TotaleCritica, SommaUtenti, TotaleUtenti, RatingAssoluto, Visualizzazioni)  
FormatoAudio(Codice, Famiglia, DataRilascio, Qualita, Bitrate)  
FormatoVideo(Codice, Famiglia, DataRilascio, Qualita, Bitrate, Risoluzione, RapportoAspetto)  
Importanza(Fattore, Utente, Valore)  
Interpretazione(Artista, Film)  
OffertaContenuto(Abbonamento, Contenuto)  
OffertaFunzionalita(Abbonamento, Funzionalita)  
Paese(Nome, InizioIP, FineIP, Latitudine, Longitudine)  
PossessoServer(Contenuto, Server)  
PremiazioneAttore(Attore, Premio)  
PremiazioneFilm(Film, Premio)  
PremiazioneRegista(Regista, Premio)  
Premio(Id, Nome, Anno, Istituzione, Descrizione, Peso, Attore, Film, Regista)  
RecensioneCritico(Critico, Film, Data, Testo, Voto)  
RecensioneUtente(Utente, Film, Data, Testo, Voto)  
RestrizioneAbbonamento(Abbonamento, Paese)  
RestrizioneContenuto(Contenuto, Paese)  
Server(Id, LarghezzaBanda, CapacitaMax, Longitudine, Latitudine, Jitter, BandaDisponibile)  
Sottotitoli(Contenuto, Lingua)  
Utente(Codice, Nome, Cognome, Email, Password, DataNascita, Nazionalita, CartaDiCredito, Abbonamento, Inizio)

### 6.2 Implementazione

Per la semantica delle tabelle e degli attributi, si rimanda alla SEZIONE 2 — PROGETTAZIONE CONCETTUALE. Tutti gli attributi, ad eccezione di quelli indicati, sono da ritenersi NOT NULL. Le coordinate geografiche sono memorizzate come DOUBLE, e convertite in base decimale.

Una breve precisazione sulla modalità di rappresentazione dell'IP *decimale con zeri*, da noi adottata di seguito, necessaria se lo si vuole memorizzare come un unico intero. La più famosa rappresentazione degli IP è quella decimale puntata, in cui a ogni byte corrisponde un numero naturale da 0 a 255, e tali numeri sono separati da punti. La trasposizione in numero intero pone il problema di distinguere tra loro i byte rappresentati da numeri con meno di 3 cifre. La nostra soluzione prevede di inserire a questo scopo degli

<sup>17</sup>Partecipando l'entità alle relationship AUDIO e CODIFICA AUDIO in maniera opzionale, potevamo scegliere, per ognuna delle relationship, se memorizzare una tabella con 2 colonne e 20mila record, da TAVOLA DEI VOLUMI, oppure aggiungere degli attributi che sarebbero stati nulli in 8mila casi. Abbiamo scelto, per ottimizzazione di spazio, la seconda opzione.

zeri. Così, ad esempio, 6.1.253.24, che sarebbe incomprensibile senza punti, diviene 006 001 253 024.

Tabella	Attributo	Codifica	Note
Abbonamento	Nome	VARCHAR(45)	Chiave primaria.
	Tariffa	INT	Costo mensile in euro.
	Durata	INT	Numero di mesi. NULLABLE.
	MaxOre	INT	Numero di ore. NULLABLE.
	EtaMinima	INT	NULLABLE.
	MaxGB	INT	Numero di Gi-gaByte. NULLABLE.
Appartenenza	Film	INT	L'Id del film. Chiave, con <i>Genere</i> .
	Genere	VARCHAR(45)	Il nome del genere. Chiave primaria, con <i>Film</i> .
Artista	Id	INT	Numero progressivo. Chiave primaria.
	Nome	VARCHAR(255)	-
	Cognome	VARCHAR(255)	-
	Pseudonimo	VARCHAR(255)	NULLABLE.
	Profilo	VARCHAR(3000)	-
	Attore	TINYINT(1)	-
	Regista	TINYINT(1)	-
CartaDiCredito	Numero	BIGINT	16 cifre. Chiave primaria.
	CVV	INT	-
	NomeIntestatario	VARCHAR(255)	-
	CognomeIntestatario	VARCHAR(255)	-

	MeseScadenza	INT	-
	AnnoScadenza	INT	-
CodificaVideo	Contenuto	INT	L'Id del Contenuto. Chiave primaria con <i>FormatoVideo</i> .
	FormatoVideo	INT	Il Codice del Formato. Chiave primaria con <i>Contenuto</i> .
Connessione	Inizio	DATETIME	Chiave primaria con <i>Dispositivo</i> .
	Dispositivo	INT	L'Id del dispositivo. Chiave primaria con <i>Inizio</i> .
	IP	BIGINT	12 cifre (IPv4). L'IP è rappresentato in notazione decimale con zeri.
	Fine	DATETIME	NULLABLE.
	Utente	INT	Il codice dell'utente.
Contenuto	Id	INT	Numero progressivo. Chiave primaria.
	Dimensione	BIGINT	In bytes.
	Lunghezza	INT	In secondi.
	Film	INT	L'Id del film.
	LinguaAudio	VARCHAR(45)	NULLABLE.
	CodificaAudio	INT	L'Id del formato. NULLABLE.
Critico	Id	INT	Numero progressivo. Chiave primaria.
	Nome	VARCHAR(255)	-
	Cognome	VARCHAR(255)	-

	Pseudonimo	VARCHAR(255)	NULLABLE.
Direzione	Film	INT	L'Id del Film. Chiave primaria con <i>Artista</i> .
	Artista	INT	L'Id dell'Artista. Chiave primaria con <i>Film</i> .
Dispositivo	Id	INT	Numero progressivo. Chiave primaria.
	Nome	VARCHAR(255)	-
	VersioneApp	VARCHAR(45)	-
	SistemaOperativo	VARCHAR(45)	-
Erogazione	Id	INT	Numero progressivo. Chiave primaria.
	Inizio	DATETIME	-
	Fine	DATETIME	NULLABLE.
	Contenuto	INT	L'Id del Contenuto.
	Server	INT	L'Id del Server.
	InizioConnessione	DATETIME	L'inizio della Connessione.
	Dispositivo	INT	L'Id del dispositivo.
Fattura	Id	INT	Numero progressivo. Chiave primaria.
	Saldo	DATE	NULLABLE.
	Scadenza	DATE	-
	Emissione	DATE	-
	Utente	INT	Il Codice dell'Utente.
	CartaDiCredito	INT	Il Numero della Carta.

	Abbonamento	VARCHAR(45)	Il Nome dell'Abbonamento.
Film	Id	INT	Numero progressivo. Chiave primaria.
	Titolo	VARCHAR(255)	-
	Descrizione	VARCHAR(3000)	-
	Anno	INT	-
	Durata	INT	In secondi.
	Paese	VARCHAR(45)	Il nome del Paese di produzione.
	SommaCritica	BIGINT	DEFAULT 0.
	TotaleCritica	INT	DEFAULT 0.
	SommaUtenti	BIGINT	DEFAULT 0.
	TotaleUtenti	BIGINT	DEFAULT 0.
	RatingAssoluto	DOUBLE	NULLABLE. DEFAULT NULL.
	Visualizzazioni	BIGINT	DEFAULT 0.
FormatoAudio	Codice	INT	Numero progressivo. Chiave primaria.
	Famiglia	VARCHAR(45)	-
	DataRilascio	DATE	-
	Qualita	VARCHAR(45)	-
	Bitrate	INT	Byte per secondo.
FormatoVideo	Codice	INT	Numero progressivo. Chiave primaria.
	Famiglia	VARCHAR(45)	-
	DataRilascio	DATE	-
	Qualita	VARCHAR(45)	-

	Bitrate	INT	Byte per secondo.
	Risoluzione	INT	-
	RapportoAspetto	VARCHAR(45)	-
Importanza	Fattore	VARCHAR(255)	Il nome del fattore. Chiave primaria, con <i>Utente</i> .
	Utente	INT	Il Codice dell'Utente. Chiave primaria, con <i>Fattore</i> .
	Valore	INT	-
Interpretazione	Artista	INT	L'Id dell'Artista. Chiave primaria con <i>Film</i> .
	Film	INT	L'Id del Film. Chiave primaria con <i>Artista</i> .
OffertaContenuto	Abbonamento	VARCHAR(45)	Il nome dell'Abbonamento. Chiave primaria, con <i>Contenuto</i> .
	Contenuto	INT	L'Id del Contenuto. Chiave primaria, con <i>Abbonamento</i> .
OffertaFunzionalita	Abbonamento	VARCHAR(45)	Il nome dell'Abbonamento. Chiave primaria, con <i>Funzionalita</i> .
	Funzionalita	VARCHAR(255)	Il nome della Funzionalità. Chiave primaria, con <i>Abbonamento</i> .
Paese	Nome	VARCHAR(45)	Chiave primaria.
	InizioIP	BIGINT	12 cifre (IPv4). L'IP è rappresentato in notazione decimale con zeri.



	FineIP	BIGINT	12 cifre (IPv4). L'IP è rappresentato in notazione decimale con zeri.
	Latitudine	DOUBLE	-
	Longitudine	DOUBLE	-
PossessoServer	Contenuto	INT	L'Id di un contenuto. Chiave primaria, con <i>Server</i> .
	Server	INT	L'Id di un Server. Chiave primaria, con <i>Contenuto</i> .
PremiazioneAttore	Attore	INT	L'Id di un artista. Chiave primaria, con <i>Premio</i> .
	Premio	INT	L'Id di un premio. Chiave primaria, con <i>Attore</i> .
PremiazioneFilm	Film	INT	L'Id di un artista. Chiave primaria, con <i>Premio</i> .
	Premio	INT	L'Id di un premio. Chiave primaria, con <i>Film</i> .
PremiazioneRegista	Regista	INT	L'Id di un artista. Chiave primaria, con <i>Premio</i> .
	Premio	INT	L'Id di un premio. Chiave primaria, con <i>regista</i> .
Premio	Id	INT	Numero progressivo. Chiave primaria.
	Nome	VARCHAR(255)	-
	Anno	INT	-
	Istituzione	VARCHAR(255)	-
	Descrizione	VARCHAR(3000)	-

	Peso	INT	-
	Attore	TINYINT(1)	-
	Film	TINYINT(1)	-
	Regista	TINYINT(1)	-
RecensioneCritico	Critico	INT	L'Id di un Critico. Chiave primaria, con <i>Film</i> .
	Film	INT	L'Id di un Film. Chiave primaria, con <i>Critico</i> .
	Data	DATE	-
	Testo	VARCHAR(3000)	-
	Voto	INT	-
RecensioneUtente	Utente	INT	L'Id di un Utente. Chiave primaria, con <i>Film</i> .
	Film	INT	L'Id di un Film. Chiave primaria, con <i>Utente</i> .
	Data	DATE	-
	Testo	VARCHAR(3000)	-
	Voto	INT	-
Restrizione Abbonamento	Abbonamento	VARCHAR(45)	Il nome di un Abbonamento. Chiave primaria, con <i>Paese</i> .
	Paese	VARCHAR(45)	Il nome di un Paese. Chiave primaria, con <i>Abbonamento</i> .
Restrizione Contenuto	Contenuto	INT	L'Id di un Contenuto. Chiave primaria, con <i>Paese</i> .

	Paese	VARCHAR(45)	Il nome di un Paese. Chiave primaria, con <i>Contenuto</i> .
Server	Id	INT	Numero progressivo. Chiave primaria.
	LarghezzaBanda	BIGINT	Misurata in byte per secondo.
	CapacitaMax	BIGINT	Misurata in byte.
	Longitudine	DOUBLE	-
	Latitudine	DOUBLE	-
	Jitter	INT	-
	BandaDisponibile	BIGINT	Misurata in byte per secondo. DEFAULT 0.
Sottotitoli	Contenuto	INT	L'Id di un Contenuto. Chiave primaria, con <i>Lingua</i> .
	Lingua	VARCHAR(45)	Il nome di una Lingua. Chiave primaria, con <i>Contenuto</i> .
Utente	Codice	INT	Numero progressivo. Chiave primaria.
	Nome	VARCHAR(255)	-
	Cognome	VARCHAR(255)	-
	Email	VARCHAR(255)	-
	Password	VARCHAR(255)	-
	DataNascita	DATE	-
	Nazionalita	VARCHAR(45)	Il Paese dell'Utente.
	CartaDiCredito	BIGINT	La Carta dell'Utente.

	Abbonamento	VARCHAR(45)	Il piano di abbonamento dell'utente. NULLABLE.
	Inizio	DATE	La data di sottoscrizione del piano di abbonamento corrente da parte dell'utente. NULLABLE.

## 6.3 Vincoli

### 6.3.1 Vincoli intrarelazionali di dominio

Valgono i seguenti:

- Gli attributi che nello schema soprastante non sono indicati con la dicitura NULLABLE hanno il vincolo NOT NULL;
- Tutti gli attributi codificati come INT hanno il vincolo di assumere valori strettamente maggiori di 0. Alcuni attributi, che presentano vincoli più restrittivi, sono indicati di seguito;
- Tutti gli attributi codificati come DATE, con l'eccezione delle date di rilascio dei formati e delle date di nascita degli utenti (le quali devono essere precedenti la current\_date al momento dell'iscrizione), devono essere successivi alla data di lancio di FilmSphere, che, considerando il tempo di vita di un anno, poniamo al 15 settembre 2022;
- Gli attributi rappresentanti coordinate geografiche devono assumere valori compresi tra -100 e +100.

**Attributi interi con particolari vincoli di dominio, estremi inclusi**

Attributo	Min	Max
CartaDiCredito(Numero)	1 000 000 000 000 000	9 999 999 999 999 999
CartaDiCredito(CVV)	100	999
CartaDiCredito(MeseScadenza)	1	12
CartaDiCredito(AnnoScadenza)	2022	-
Connessione(IP)	0	255255255255
Fattura(CartaDiCredito)	1 000 000 000 000 000	9 999 999 999 999 999
Film(Anno)	1900	YEAR(CURRENT_DATE)
Importanza(Valore)	1	100
Paese(InizioIP)	0	255255255255

Paese(FineIP)	0	255255255255
Premio(Peso)	0	100
RecensioneCritico(Voto)	1	5
RecensioneUtente(Voto)	1	5
Server(Jitter)	1	10
Utente(CartaDiCredito)	1 000 000 000 000 000	9 999 999 999 999

### 6.3.2 Vincoli intrarelazionali di n-upla

Valgono i seguenti:

- Artista(Attore) e Artista(Regista) non possono essere entrambi nulli;
- La data definita dagli attributi CartaDiCredito(Mese) e CartaDiCredito(Anno) deve essere futura;
- Paese(InizioIP) deve essere minore di Paese(FineIP);
- Premio(Attore), Premio(Film), Premio(Regista) non possono essere tutti nulli.

### 6.3.3 Vincoli interrelazionali di integrità referenziale

Valgono i seguenti:

L'attributo	Presenta un v.i.r. verso
Appartenenza(Film)	Film(Id)
CodificaVideo(Contenuto)	Contenuto(Id)
CodificaVideo(FormatoVideo)	FormatoVideo(Codice)
Connessione(Dispositivo)	Dispositivo(Id)
Connessione(Utente)	Utente(Codice)
Contenuto(Film)	Film(Id)
Contenuto(CodificaAudio)	FormatoAudio(Codice)
Direzione(Film)	Film(Id)
Direzione(Artista)	Artista(Id)
Erogazione(Contenuto)	Contenuto(Id)

Erogazione(Server)	Server(Id)
Erogazione(InizioConnessione)	Connessione(Inizio)
Erogazione(Dispositivo)	Dispositivo(Id)
Fattura(Utente)	Utente(Codice)
Fattura(CartaDiCredito)	CartaDiCredito(Numero)
Fattura(Abbonamento)	Abbonamento(Nome)
Film(Paese)	Paese(Nome)
Importanza(Utente)	Utente(Codice)
Interpretazione(Artista)	Artista(Id)
Interpretazione(Film)	Film(Id)
OffertaContenuto(Abbonamento)	Abbonamento(Nome)
OffertaContenuto(Contenuto)	Contenuto(Id)
OffertaFunzionalita(Abbonamento)	Abbonamento(Nome)
PossessoServer(Contenuto)	Contenuto(Id)
PossessoServer(Server)	Server(Id)
PremiazioneAttore(Artista)	Artista(Id)
PremiazioneAttore(Premio)	Premio(Id)
PremiazioneFilm(Film)	Film(Id)
PremiazioneFilm(Premio)	Premio(Id)
PremiazioneRegista(Artista)	Artista(Id)
PremiazioneRegista(Premio)	Premio(Id)
RecensioneCritico(Critico)	Critico(Id)
RecensioneCritico(Film)	Film(Id)
RecensioneUtente(Utente)	Utente(Codice)
RecensioneUtente(Film)	Film(Id)
RestrizioneAbbonamento(Abbonamento)	Abbonamento(Nome)

RestrizioneAbbonamento(Paese)	Paese(Nome)
RestrizioneContenuto(Contenuto)	Contenuto(Id)
RestrizioneContenuto(Paese)	Paese(Nome)
Sottotitoli(Contenuto)	Contenuto(Id)
Utente(Nazionalita)	Paese(Nome)
Utente(CartaDiCredito)	CartaDiCredito(Numero)
Utente(Abbonamento)	Abbonamento(Nome)

### 6.3.4 Altri vincoli interrelazionali

Valgono i seguenti:

- Un Artista può comparire nella tabella *Interpretazione* solo se il suo attributo *Attore* vale 1;
- Un Artista può comparire nella tabella *Direzione* solo se il suo attributo *Regista* vale 1;
- Un Artista può comparire nella tabella *PremiazioneAttore* solo se il suo attributo *Attore* vale 1;
- Un Premio può comparire nella tabella *PremiazioneAttore* solo se il suo attributo *Attore* vale 1;
- Un Premio può comparire nella tabella *PremiazioneFilm* solo se il suo attributo *Film* vale 1;
- Un Artista può comparire nella tabella *PremiazioneRegista* solo se il suo attributo *Regista* vale 1;
- Un Premio può comparire nella tabella *PremiazioneRegista* solo se il suo attributo *Regista* vale 1;
- Un contenuto può avere l'attributo *CodificaAudio* nullo se e solo se compare nella tabella *CodificaVideo*;
- Un contenuto compare nella tabella *Sottotitoli* se e solo se il suo attributo *LinguaAudio* è nullo.

## 7 — Analisi delle dipendenze funzionali e normalizzazione

La maggior parte delle relazioni del database presenta un'unica chiave, e, come unica dipendenza funzionale non banale, quella in cui la chiave implica tutti gli altri attributi della n-upla.

È il caso delle seguenti relazioni:

Abbonamento, Appartenenza, Artista<sup>18</sup>, CartaDiCredito, CodificaVideo, Connessione, Contenuto, Critico, Direzione, Dispositivo, Film, FormatoAudio, FormatoVideo, Importanza, Interpretazione, OffertaContenuto, OffertaFunzionalita, PossessoServer, PremiazioneAttore, PremiazioneFilm, PremiazioneRegista, Premio, RecensioneCritico, RecensioneUtente, RestrizioneAbbonamento, RestrizioneContenuto, Sottotitoli.

Esse sono, per definizione, in Forma Normale di Boyce-Codd.

<sup>18</sup>Questa relazione, come *Critico*, *Film* e *Premio*, necessita di ulteriori chiarimenti. Per garantire alla nostra Base di Dati la flessibilità richiesta da una società multinazionale, quale è FilmSphere, abbiamo infatti scelto di non considerare, in queste relazioni, dipendenze funzionali diverse da quelle della chiave verso l'intera n-upla. In questo modo, ad esempio, non ci precludiamo a priori la possibilità di omonimie, riservando ad altre aree dell'azienda l'onere di garantire all'utenza una comunicazione chiara ed efficace.

Analizziamo ora, caso per caso, le relazioni restanti:

Nella tabella **Erogazione** sono presenti tre dipendenze funzionali:

$\text{Id} \rightarrow \text{intera n-upla}$

$\text{Inizio}, \text{InizioConnessione}, \text{Dispositivo} \rightarrow \text{intera n-upla}$

$\text{Fine}, \text{InizioConnessione}, \text{Dispositivo} \rightarrow \text{intera n-upla}$

Ipotizziamo infatti, come è ovvio, che una connessione non possa essere coinvolta contemporaneamente in due erogazioni.

Ogni implicante costituisce una chiave della relazione; di conseguenza **Erogazione** è in BCNF.

Nella tabella **Fattura** sono presenti tre dipendenze funzionali:

$\text{Id} \rightarrow \text{intera n-upla}$

$\text{Emissione}, \text{Utente} \rightarrow \text{intera n-upla}$

$\text{Emissione} \rightarrow \text{Scadenza}$

Poiché, infatti, la prima fattura relativa a un piano di abbonamento viene emessa a un mese dalla sottoscrizione dello stesso, e un utente non può sottoscrivere due piani di abbonamento contemporaneamente, non è in alcun modo possibile che vengano emesse, nello stesso giorno, due fatture relative allo stesso utente.

La terza dipendenza funzionale è motivata dal fatto che, nell'OPERAZIONE 8, la scadenza di una fattura è sempre posta a 30 giorni dall'emissione della stessa. La norma, in questo caso, vorrebbe che si attuasse una decomposizione della relazione secondo il relativo algoritmo, producendo le tabelle **Fattura**(Id, Saldo, Emissione, Utente, CartaDiCredito, Abbonamento) e **NuovaTabella**(Emissione, Scadenza). Abbiamo tuttavia ritenuto che anche questa soluzione fosse ridondante, dal momento che, senza alcuna informazione aggiuntiva, dalla data di emissione è sempre possibile ricavare la data di scadenza. Abbiamo pertanto deciso di rimuovere l'attributo **Scadenza** dalla relazione **Fattura**.

Nella tabella **Paese** sono presenti quattro dipendenze funzionali:

$\text{Nome} \rightarrow \text{intera n-upla}$

$\text{InizioIP} \rightarrow \text{intera n-upla}$

$\text{FineIP} \rightarrow \text{intera n-upla}$

$\text{Latitudine}, \text{Longitudine} \rightarrow \text{intera n-upla}$

Ogni implicante costituisce una chiave della relazione; di conseguenza **Paese** è in BCNF.

Nella tabella **Server** sono presenti due dipendenze funzionali:

$\text{Id} \rightarrow \text{intera n-upla}$

$\text{Longitudine}, \text{Latitudine} \rightarrow \text{intera n-upla}$

Ogni implicante costituisce una chiave della relazione; di conseguenza **Server** è in BCNF.

Nella tabella **Utente** sono presenti due dipendenze funzionali:

$\text{Codice} \rightarrow \text{intera n-upla}$

$\text{Email} \rightarrow \text{intera n-upla}$

Ogni implicante costituisce una chiave della relazione; di conseguenza **Utente** è in BCNF.

Tutte le tabelle sono quindi in BCNF.



## 8 — Data Analytics

### 8.1 Classifiche

Tra i piani su cui competono i servizi di streaming spicca la scelta dei contenuti: le aziende sono continuamente chiamate a selezionare prodotti e a investire su di essi. Nel medio periodo, la scelta dei contenuti ha un impatto notevolissimo sull'identità del marchio e porta alla formazione, talvolta, di una vera e propria filosofia aziendale.

L'operazione, che è spesso svolta su base trimestrale, è resa complessa da 3 fattori:

- L'immensa mole di film: secondo IMDB, dal 1900 a oggi sono stati prodotti in media 2577 film l'anno;
- La variabilità dei gusti tra i vari Paesi del mondo: lo stesso film potrebbe essere un ottimo investimento in un'area geografica e un pessimo investimento in un'altra. Un simile divario si registra anche tra le varie classi sociali;
- La difficoltà nell'avere un riscontro da parte degli utenti, e in particolare nell'individuare i contenuti preferiti dal pubblico.

Un ragionamento simile può essere fatto in merito ai formati: utilizzare contenuti codificati in un determinato formato significa investire in quel formato, credere che sarà utilizzato in futuro e che garantirà agli utenti un'esperienza di streaming soddisfacente. Inoltre, come i film, i formati sono moltissimi, e diffusi in modo eterogeneo nei vari Paesi.

La funzionalità che proponiamo mira a supportare l'azienda nell'individuazione di formati e contenuti su cui investire tramite l'analisi delle visualizzazioni da parte degli utenti. Tale analisi viene svolta in maniera separata sui vari segmenti dell'utenza di FilmSphere, in termini di nazionalità e piano di abbonamento. In questo modo i manager della società hanno a loro disposizione un rapporto puntuale sull'utilizzo che gli utenti fanno di film e formati, e tale rapporto tiene conto dell'area geografica e, tramite il piano di abbonamento, dell'estrazione sociale degli utenti.

La procedura, che non richiede parametri in input, restituisce, per ogni Paese e piano di abbonamento, la classifica dei film più visualizzati corredata dai secondi di visualizzazione. Inoltre, per ogni film, piano di abbonamento e area geografica, la procedura restituisce la classifica dei formati per percentuale di fruizione del film da parte di utenti di quella area geografica con quell'abbonamento, corredata dalla relativa percentuale di fruizione. I dati restituiti sono relativi all'ultimo trimestre.

La procedura da noi proposta, tuttavia, non fornisce da sola all'azienda i dati sufficienti a operare una scelta consapevole, in quanto non può fornire dati su film che non siano già presenti su FilmSphere. In altre parole, la procedura è utile a individuare formati e film da dismettere, o a isolare acquisizioni di successo su cui puntare in futuro: non può supportare l'azienda nell'analisi di nuovi film da aggiungere al portale. L'azienda dovrà supplire a questa mancanza attraverso adeguati strumenti di analisi del mercato, per i quali, eventualmente, potrà affidarsi a società terze specializzate.

### 8.2 Bilanciamento del carico

Un'altra esigenza di programmazione aziendale molto sentita concerne la Content Delivery Network, che costituisce il tramite tra i contenuti di FilmSphere e l'utente finale. Particolarmente importante è la prevenzione dei sovraccarichi, che pregiudicano significativamente la qualità di fruizione dei contenuti da parte degli utenti e, alla lunga, possono portare gli stessi a perdere fiducia nel servizio. La prevenzione dei sovraccarichi può essere posta in essere su due fronti:

- L'individuazione delle aree della rete che presentano un maggiore rischio di sovraccarico e l'espansione, o il consolidamento, della rete in tali aree;
- Lo sviluppo di piani di allocazione dei contenuti che permettano di ottimizzare le prestazioni dei server esistenti.

La procedura da noi proposta utilizza entrambi gli approcci. In primo luogo, a una variabile passata in input viene assegnato l'indice  $\sigma$ , che quantifica il rischio di sovraccarico del server. Tale indice viene calcolato a

partire dalle percentuali di banda utilizzata dal server nei mesi del trimestre<sup>19</sup>, così determinate:

$$\beta_i = \frac{100 * \sum_{i=1}^n \dim(c_i)}{B_{tot} * 24 * 3600 * 30}$$

dove  $B_{tot}$  è la banda totale del server,  $n$  è il numero di contenuti al momento in erogazione da parte del server nel mese,  $\dim(c_i)$  e  $dur(c_i)$  restituiscono, rispettivamente, dimensione (in byte), e durata (in secondi) di un contenuto  $i$ .

$$\sigma = (\beta_3 + \frac{(100-\beta_3)*(\beta_3-\beta_2)}{2} + \frac{(100-\beta_3)*(\beta_3-\beta_1)}{2}) * \frac{1}{2} + 50$$

Dove numero 1 si riferisce al primo mese, del trimestre, il numero 2 al secondo e il numero 3 al terzo. Il coefficiente  $\sigma$  tiene conto della percentuale di banda impiegata dal server nell'ultimo mese del trimestre e dell'incremento di tale percentuale rispetto al primo e al secondo mese. Così concepito,  $\sigma$  è un numero compreso tra -100 e +100; a seguito della divisione per 2 e la somma di un valore 50, esso diventa compreso tra 0 e 100. Lo studio dell'incremento percentuale è, a nostro avviso, molto importante, per prevedere nella nostra metrica i dati futuri sulla base della tendenza pregressa.

In secondo luogo, vengono individuati  $n$  spostamenti strategici di contenuti verso altri server in caso di sovraccarico. La tabella con tali spostamenti viene stampata a video. Nel caso in cui  $n$  fosse maggiore del numero di contenuti posseduti dal server,  $n$  viene troncato a tale numero. A rendere uno spostamento strategico concorrono i seguenti fattori:

- Il contenuto è molto visualizzato dagli utenti connessi al server target, per cui un malfunzionamento nel server target avrebbe grande impatto;
- La dimensione del contenuto è elevata, per cui la sua erogazione occupa molta banda;
- La differenza di banda disponibile tra il server di destinazione e quello target è elevata;
- Il server di destinazione è geograficamente vicino al server target.

Naturalmente, il server di destinazione non deve possedere, al momento attuale, il contenuto oggetto dello spostamento.

Tali fattori sono compendati nel coefficiente  $\eta$ , di seguito proposto, che viene calcolato per ogni contenuto e per ogni server che non possieda tale contenuto.

$$\eta = \frac{vis(\gamma) * \dim(\gamma) * ((B_{tot,d} * 30 * 24 * 3600 - \sum_{i=1}^n \dim(c_i)) - (B_{tot,t} * 30 * 24 * 3600 - \sum_{i=1}^n \dim(c_i)))}{dist(S_d, S_t)}$$

Dove  $vis(\gamma)$  restituisce il numero di visualizzazioni del contenuto negli ultimi 30 giorni in erogazioni che coinvolgano il server target,  $\dim(\gamma)$  restituisce la dimensione, in MegaByte, del contenuto, il successivo fattore esprime la differenza di banda disponibile media tra i server negli ultimi 30 giorni, in TeraByte e  $dist(S_1, S_2)$  restituisce la distanza geografica, in metri, tra i server. Gli  $n$  spostamenti con il coefficiente  $\eta$  maggiore vengono suggeriti dalla procedura.

Il bilanciamento del carico è un'operazione complessa con cui la nostra metrica si interfaccia in modo rudimentale. Tra le maggiori criticità, spiccano il non tenere conto di tendenze cicliche giornaliere e annuali (e.g., le visualizzazioni dei film di Natale), delle differenze nei momenti di carico e scarico tra i server, della saturazione delle reti Internet nell'area antistante i nostri server, e in generale di algoritmi di scheduling real-time dell'erogazione dei contenuti da parte dei server.

### 8.3 Fruizione media dei vincoli dell'abbonamento

Un'altra sfida importante per un'azienda di streaming risiede nella formulazione di un'offerta commerciale attraente e competitiva. I piani di abbonamento di una società come FilmSphere devono essere preferiti,

<sup>19</sup>Dal momento che questa operazione può essere eseguita in ogni giorno dell'anno, con *trimestre* si designano gli ultimi 90 giorni e con *mese* un insieme di 30 giorni consecutivi.

dagli utenti, a quelli dei competitor, e d'altra parte devono assicurare margini di profitto sufficienti. Particolarmente importante, nell'ideazione di nuovi piani di abbonamento e nella revisione di quelli attuali, si rivela l'analisi dell'interazione degli utenti con le offerte che hanno sottoscritto, e in particolare con i vincoli posti dalle stesse. Nella concezione di FilmSphere abbiamo immaginato che tali vincoli riguardino la dimensione e la durata dei contenuti fruibili mensilmente. Tale analisi diviene tanto più rigorosa e utile quanto più è mirata, da un punto di vista geografico e generazionale.

Proponiamo quindi una funzionalità analytics che, segmentando la popolazione per Paese, fascia d'età e abbonamento sottoscritto, restituisca le percentuali medie, nel mese e nell'anno dati in input, della fruizione di contenuti da parte degli utenti rispetto ai vincoli dell'abbonamento, in termini di byte e di secondi.

Tale operazione può, nel medio e lungo periodo, costituire a nostro avviso uno strumento fondamentale nella programmazione dell'azienda. A titolo esemplificativo, sarà ideale un abbonamento che permetta di fruire moltissimi contenuti, se ad esso corrisponde una bassa fruizione effettiva, in quanto tale offerta sarà massimamente allettante per un potenziale utente e inciderà poco sui costi dell'azienda. Si riporta di seguito un prospetto con le fasce d'età, ripreso dall'OPERAZIONE 5.

Fascia	Anni
Teen	13 - 18
Young	19 - 35
Middle-Aged	36 - 65
Senior	65 +

La nostra operazione, tuttavia, non riesce a fornire da sola un quadro completo del successo dei piani d'abbonamento di FilmSphere, poiché nulla dice sui piani che non presentano vincoli. L'amministrazione di FilmSphere avrà quindi bisogno di integrarla con funzionalità ausiliarie.

## 8.4 Lingue per tempo di fruizione, come audio e come sottotitoli

Proponiamo un'ultima, semplice funzionalità di analytics business-oriented, che studi la fruizione di contenuti nelle varie lingue da parte degli utenti, e l'incremento trimestrale di tali dati. Sulla concezione di trimestre, vale quanto detto nella NOTA 19. I dati sulla fruizione di contenuti in ogni lingua sono essenziali per determinare quali contenuti acquisire e caricare su FilmSphere: se, ad esempio, le ore di riproduzione di contenuti in una lingua sono minime, si può considerare di rimuovere tale lingua da FilmSphere. Al contrario, se le statistiche di riproduzione di una lingua registrano una forte crescita trimestrale, è possibile valutare l'acquisizione di un gran numero di contenuti in tale lingua.

## 9 — Popolamento del database

Per valorizzare al meglio le funzionalità del nostro progetto senza compromettere l'efficienza dei test, abbiamo popolato le varie tabelle del database riducendone i volumi di una scala variabile. Riportiamo un prospetto esplicativo.

Concetto	VolumeTeorico	VolumeDB
----------	---------------	----------

Abbonamento	5	5
Appartenenza	750	15
Artista	2960	30, di cui 21 solo attori, 6 solo registi e 3 entrambi.
Audio	20mila	400
Carta di Credito	120mila	120
Codifica Video	8mila	160
Connessione	47.45 milioni	950
Contenuto	28mila	560
Critico	30	$4 + 1^{20}$
Direzione	600	12
Dispositivo	200mila	200
Erogazione	94.9 milioni	1900
Fattore	7	7
Fattura	780mila	780
Film	500	10
Formato Audio	30	3
Formato Video	60	6
Funzionalità Aggiuntiva	5	5
Genere	10	4
Importanza	600mila	600
Interpretazione	5600	42
Lingua	8	8
Offerta Contenuto	110mila	2200

---

<sup>20</sup>Aggiungiamo un quinto critico, a cui non sono relative recensioni, per meglio mostrare le funzionalità dell'OPERAZIONE 8.

Offerta Funzionalità	10	10
Paese	40	24
Possesso Server	100mila	1400
Premiazione Attore	1120	10
Premiazione Film	750	15
Premiazione Regista	320	6
Premio	2mila	20
Recensione Critico	2mila	40
Recensione Utente	300mila	300
Restrizione Abbonamento	15	10
Restrizione Contenuto	184800	2310
Server	10	7
Sottotitoli	40mila	800
Utente	100mila	100

Riportiamo di seguito, inoltre, una possibile formulazione dei piani di abbonamento, a cui ci siamo attenuti nel popolamento del database.

Nome	Tariffa	Età Minima	Max GB	Max Ore	Funzionalità aggiuntive	Numero di contenuti visualizzabili
Basic	5	-	400	300	-	200
Premium	8	-	700	600	-	400
Deluxe	13	-	1200	1000	No pubblicità, Accessi simultanei dallo stesso profilo	500

Pro	17	14 <sup>21</sup>	-	-	No pubblicità, Accessi simultanei dallo stesso profilo, Visione dei film con un amico.	530
Ultimate	22	14	-	-	No pubblicità, Accessi simultanei dallo stesso profilo, Visione dei film con un amico, Download di contenuti offline, Generazione automatica dei sottotitoli nelle lingue non supportate.	560

Dal momento che la funzionalità "Download di contenuti offline" non è ancora disponibile in tutti i Paesi, il piano Ultimate è per ora soggetto a restrizioni in 15 Paesi. Tutti gli abbonamenti tranne Ultimate, che non la ha, hanno durata 1 anno.

---

<sup>21</sup>Qui e di seguito, il limite d'età è dovuto a ragioni legali connesse alla trasmissione del volto e della voce dell'utente nella funzionalità "Visione dei film con un amico"