

Elaborato per il corso di  
“Basi di Dati”

*Progetto di una base di dati per la gestione di dati di un'applicazione*

Pietro Ventrucchi  
pietro.ventrucchi@studio.unibo.it  
00001031205

2 settembre 2023

# Indice

<b>1</b>	<b>Analisi dei requisiti</b>	<b>2</b>
1.1	Intervista . . . . .	2
1.2	Estrazione dei concetti principali . . . . .	3
<b>2</b>	<b>Progettazione concettuale</b>	<b>5</b>
2.1	Schema scheletro . . . . .	5
2.2	Raffinamenti proposti . . . . .	6
2.3	Schema concettuale finale . . . . .	9
<b>3</b>	<b>Progettazione Logica</b>	<b>12</b>
3.1	Stima dei volumi . . . . .	12
3.2	Descrizione delle operazioni principali e stima della loro frequenza . . . . .	14
3.3	Tabelle degli accessi e schemi di navigazione . . . . .	15
3.4	Analisi delle ridondanze . . . . .	20
3.5	Raffinamento dello schema . . . . .	22

# Analisi dei requisiti

Si vuole creare una base di dati a supporto dell'applicazione *SportConnected* per gestire le attività svolte dagli utenti. La base di dati immagazzina quindi informazioni su utenti e relativa attività al fine di poter fornire una consultazione comoda per visualizzare i progressi del proprio allenamento o degli altri.

## 1.1 Intervista

Una persona si registra e viene identificata univocamente nel sistema da un codice, anche un nome, cognome, data di nascita; opzionalmente una foto profilo, una biografia, città e provincia, peso, altezza. Si può anche indicare il tipo di attrezzatura utilizzata (tipologia di bici o marca di scarpe). Il tipo di sport, la durata dell'attività, il genere, gli anni, il peso e l'altezza possono servire per calcolare le calorie spese, la potenza o altri dati utili. Un utente può seguire o essere seguito da altri utenti (potrebbero essere suggeriti in base ad amicizie comuni). Ogni utente ha un registro delle proprie attività. Possono anche essere salvati dei percorsi o dei segmenti, relativi ad attività di altri utenti, o proprie attività i quali vengono salvati sotto forma di file .gpx.

Un percorso rappresenta l'intero itinerario dell'attività e può contenere o meno dei segmenti, mentre un segmento è una sezione con determinate caratteristiche, quali salita, discesa, tratto difficile, acque libere...Ogni tratto completato ha un tempo di percorrenza associato. Le attività hanno i propri segmenti e le proprie classifiche.

Possono inoltre essere creati segmenti dagli utenti, a condizione che il segmento sia presente nel percorso svolto e che non ne siano presenti altri simili, potrà essere impostato come pubblico o privato. Possono essere assegnati dei riconoscimenti all'utente in base ai tempi raggiunti sui segmenti, titoli come: più veloce di tutti, top 10, record personale...

Le attività possono essere caricate, decidendo la visibilità, il titolo, una descrizione, il tipo di sport, un percorso, un giorno, distanza totale, dislivello, velocità media, tempo, calorie bruciate, altitudine...Vengono visualizzati anche i relativi segmenti del percorso. Possono poi essere commentate da altri utenti in grado di assegnare anche un apprezzamento.

## 1.2 Estrazione dei concetti principali

### Glossario dei termini

Si estraggono dall'intervista i concetti principali, fornendone una breve descrizione, eventuali sinonimi e relazioni con altri concetti.

Termine	Descrizione	Sinonimi	Relazioni
Utente	Persona che si registra all'applicazione. Può caricare le sue attività o salvare percorsi e segmenti fatti da altri utenti. Può interagire con altri utenti.	Persona	Attrezzatura, Attività, Commento, Percorso, Segmento
Attrezzatura	Materiale utilizzato da un utente per svolgere le proprie attività.	Bici, Scarpe	Utente
Attività	L'insieme di tutte le informazioni dell'attività sportiva svolta da un utente.	Sport	Utente, Percorso, Commento, Sport
Sport	Tipologia di attività effettuata. Contraddistingue gli sport applicabili ad un segmento.	Attività	Attività, Segmento
Percorso	Intera tratta percorsa in un'attività.	Itinerario	Attività, Segmento, Utente
Segmento	Sezione del percorso con determinate caratteristiche. Possono essere creati dagli utenti.	Tratto	Percorso, Utente, Riconoscimento
Riconoscimento	Titolo assegnato ad un utente sul segmento in base al tempo di impiegato.	Titolo	Utente, Segmento
Commento	Messaggio lasciato da utenti inerente all'attività svolta da altri utenti.		Utente, Attività

### Ristrutturazione dei requisiti

A seguito dell'analisi dei requisiti e all'estrazione dei concetti principali si procede eliminando omonimie e riformulando la richiesta in maniera che risulti più chiara e fruibile per la realizzazione della base di dati.

Per ogni **utente** vengono salvati un codice, nome, cognome, data di nascita, sesso e opzionalmente una foto profilo, biografia, città e provincia, peso, altezza, **attrezzatura** utilizzata. Un utente può seguire o essere seguito da altri utenti. Ad ogni utente vengono associate le relative **attività** e può salvare **percorsi** o **segmenti** di altri utenti. Le informazioni sull'utente unite a quelle dell'**attività** potrebbero essere utilizzate per calcolare calorie bruciate, potenza e altri dati relativi all'allenamento.

L'**attrezzatura** è indicabile scegliendo tipologia di bici o marca di scarpe. Ogni utente può specificare un nickname, non può utilizzare più volte lo stesso nickname, e un modello a piacere (inseribile a piacere, non esiste un elenco di modelli da cui scegliere).

Le **attività** devono specificare titolo, visibilità, **percorso**, giorno, distanza totale, tempo, velocità media...In base allo **sport** effettuato hanno più o meno dati rilevanti. Le attività rese pubbliche possono ricevere **commenti** e like da altri **utenti**.

Lo **sport** relativo ad un'attività o ad un segmento.

Il **percorso** è rappresentato da una traccia .gpx, una descrizione, un nome e da possibili **segmenti**.

Un **segmento** è rappresentato da una traccia .gpx, una determinata caratteristica, un tempo di percorrenza associato in attività, visibilità pubblica o privata e un nome. Possono essere percorsi più volte durante un singolo percorso. I segmenti sono relativi al tipo di **sport** e possono essere creati dagli **utenti** se presenti nel **percorso** caricato in un'attività e non ne sono presenti di simili.

Un **riconoscimento** può essere assegnato ad un **utente** in base al tempo impiegato su un determinato **segmento**.

I **commenti** sono riferiti ad un'attività e scritti da un **utente**, possono ricevere like.

# Progettazione concettuale

## 2.1 Schema scheletro

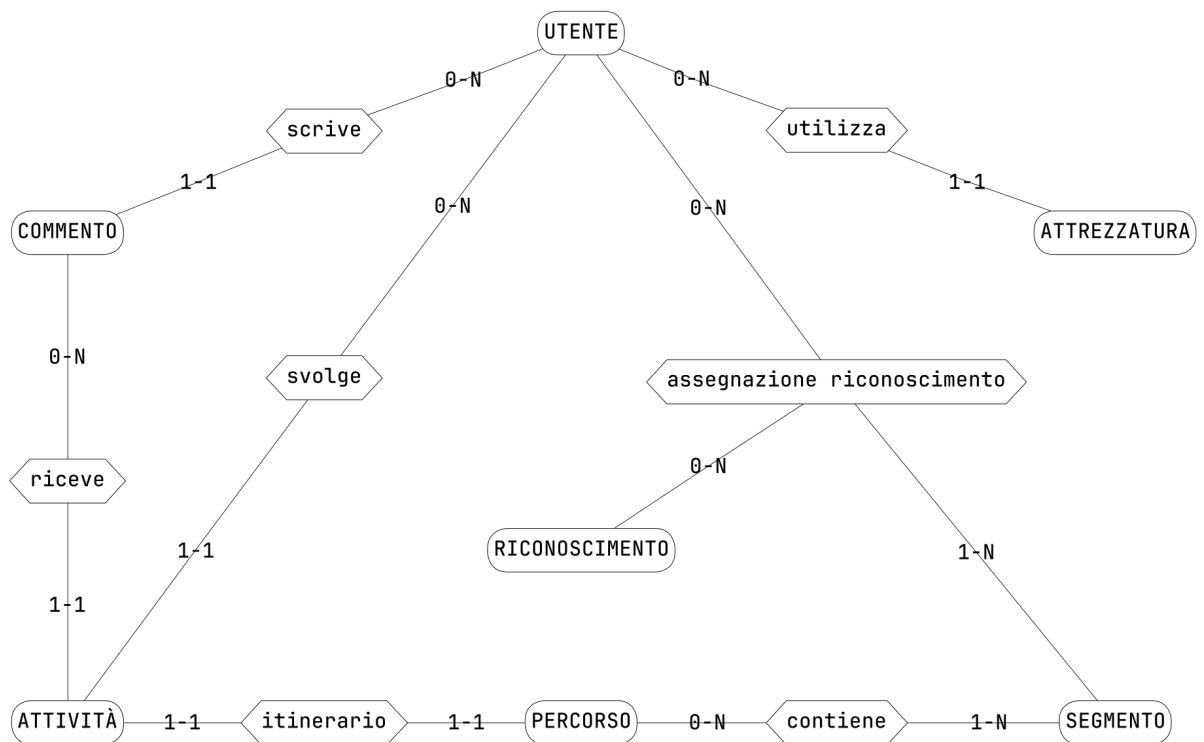


Figura 2.1: *Schema scheletro con le principali entità.*

## 2.2 Raffinamenti proposti

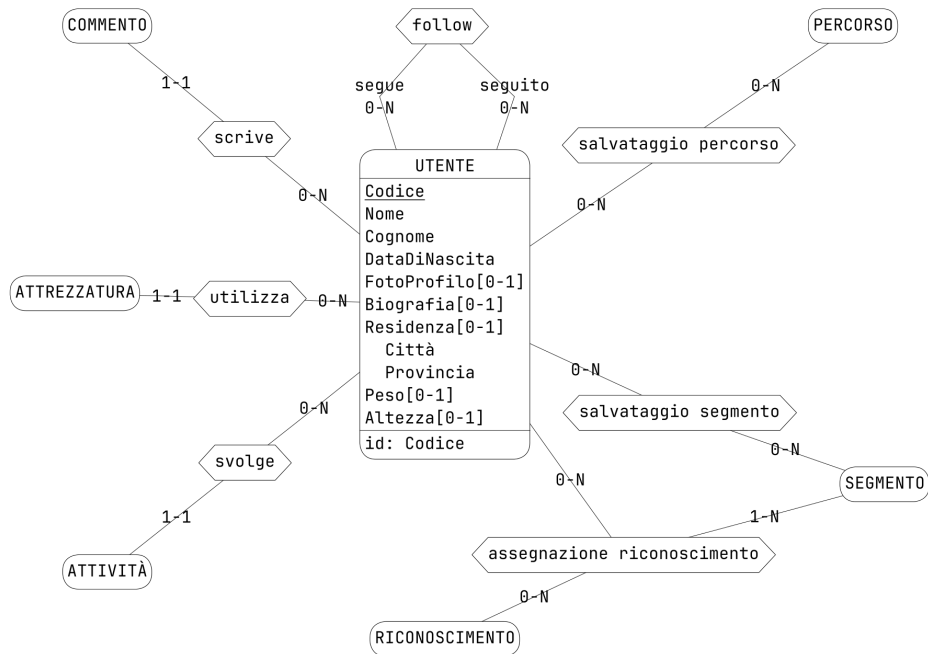


Figura 2.2: *Schema raffinamento utente.*

Per un utente, identificato da un codice univoco, è rappresentato il nome, il cognome, la data di nascita e opzionalmente il percorso ad una foto profilo caricata, una biografia, la residenza composta da città e provincia, il peso e l'altezza. La auto-riferenziazione in utente serve per esprimere il concetto di utenti seguiti e che seguono altri utenti.

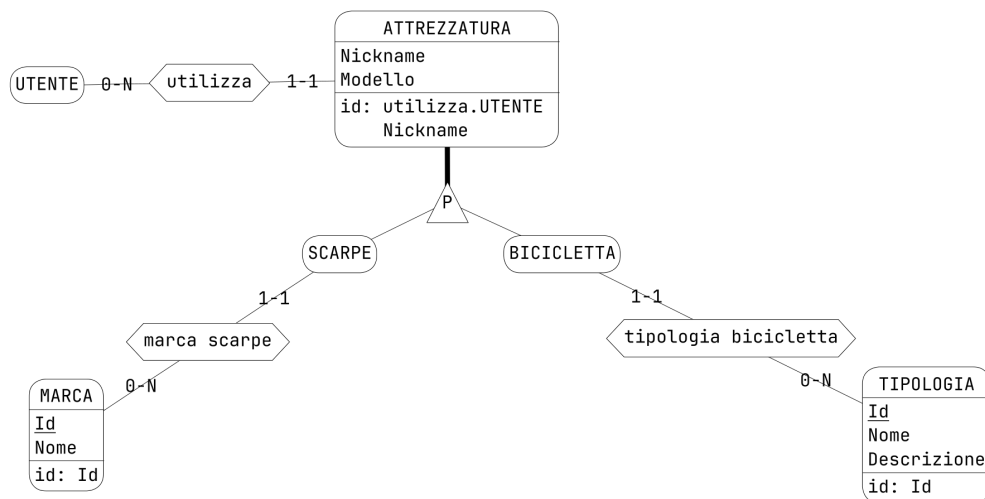


Figura 2.3: *Schema raffinamento attrezzatura.*

L'entità attrezzatura identificata da un nickname è univoca per ogni utente e rappresenta una generalizzazione delle due tipologie di attrezzatura disponibili: bicicletta e scarpe, viene quindi utilizzata la specializzazione, tramite generalizzazione totale ed esclusiva in queste due sottoclassi. Devono esistere nel DB delle entità che contengano tutte le tipologie di bici o le marche di scarpe associabili a bici e scarpe. L'attributo modello viene lasciato ad attrezzatura in quanto sia che si tratti di scarpe che di bici andrà specificato e sarà inserito a piacere dall'utente.

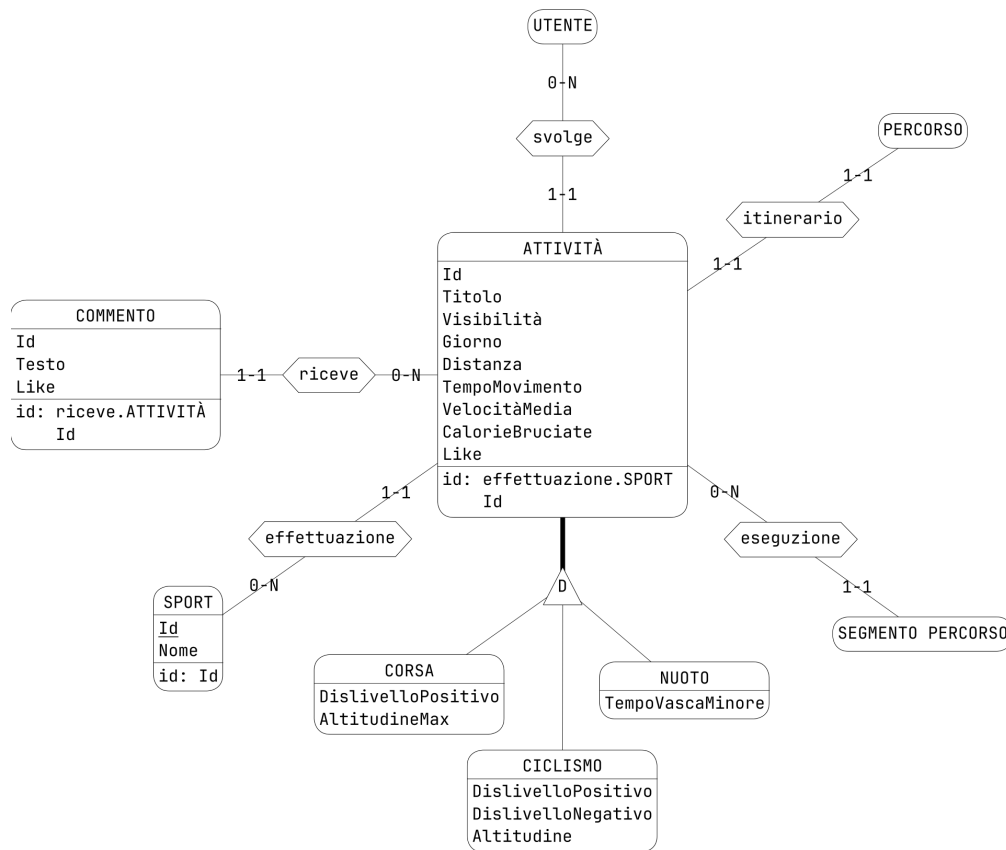


Figura 2.4: Schema raffinamento attività.

Ogni attività che viene inserita deve essere obbligatoriamente associata anche ad uno sport, il quale può avere delle determinate caratteristiche che sono specificate dalle entità figlie di attività, rimane inespresso il vincolo per cui gli sport dell'entità sport siano associati alle entità figlie di attività. La generalizzazione è parziale ed esclusiva, perchè possano essere aggiunti facilmente anche altri sport. Viene salvato anche un percorso, il quale è univoco per ogni attività e contiene o meno dei segmenti. Possono ricevere commenti scritti da utenti ed ogni commento ha un codice univoco riferito all'attività dato il grande numero di commenti che potrebbero essere scritti il numero potrebbe crescere troppo ed è preferibile tenerlo basso.



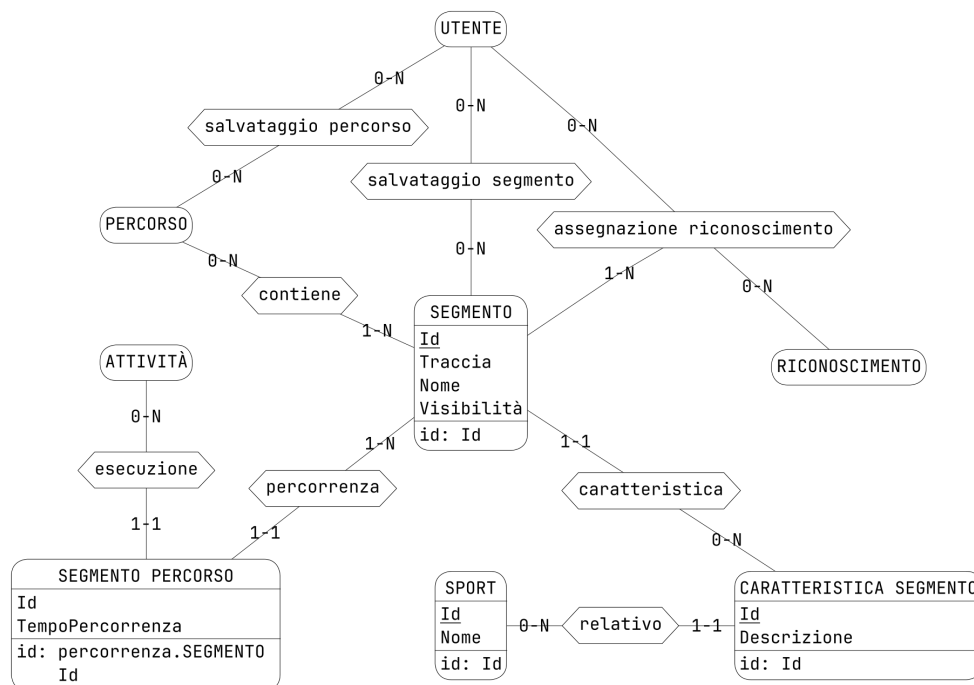


Figura 2.5: Schema raffinamento segmento.

Il segmento è un'entità che contiene tutti i segmenti esistenti, mentre il tempo di percorrenza del segmento durante l'attività viene registro tramite segmento percorso. Ogni segmento ha una caratteristica che si riferisce ad uno sport. Rimane inespresso il vincolo per il quale i segmenti facenti parte del percorso sono gli stessi di segmento percorso collegato ad attività.

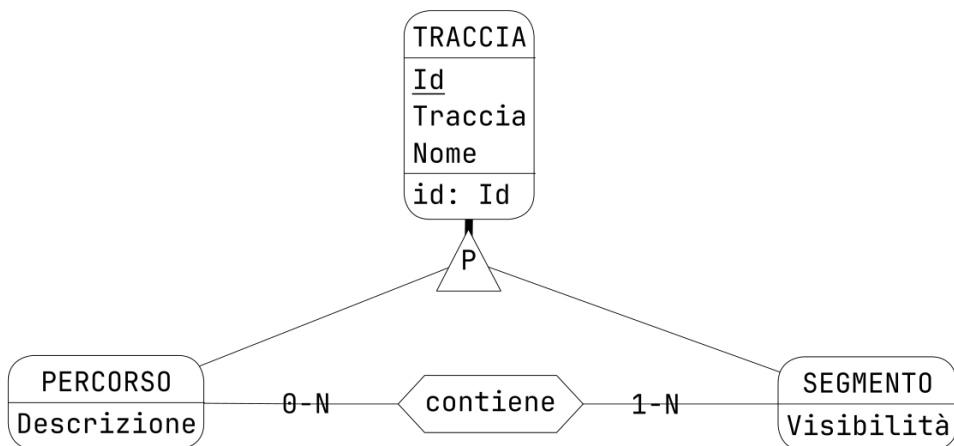


Figura 2.6: Schema raffinamento percorso e segmento.

I percorsi e i segmenti hanno entrambi una traccia, un nome e si riferiscono a tratti di strada od altro percorsi durante un'attività, per questo motivo vengono generalizzati dall'en-

tità traccia, con generalizzazione totale ed esclusiva, dove l'attributo traccia rappresenta il percorso in memoria nel quale viene salvata la traccia .gpx.

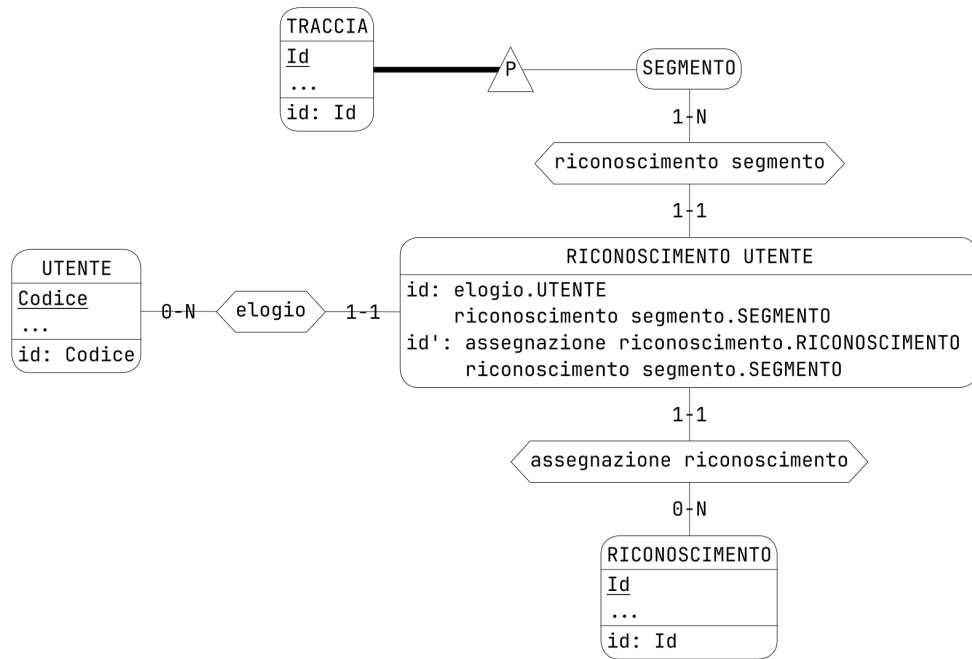
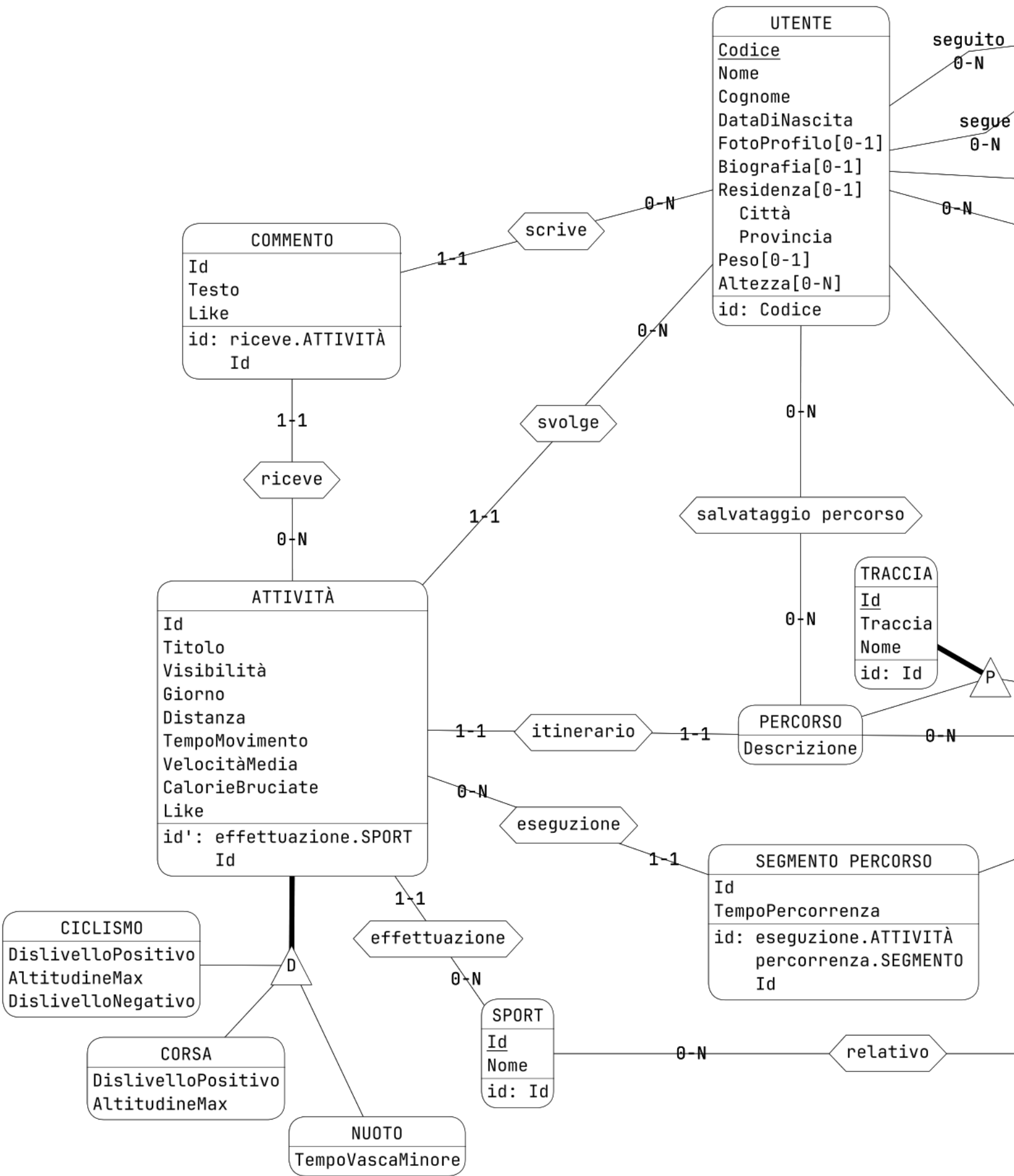


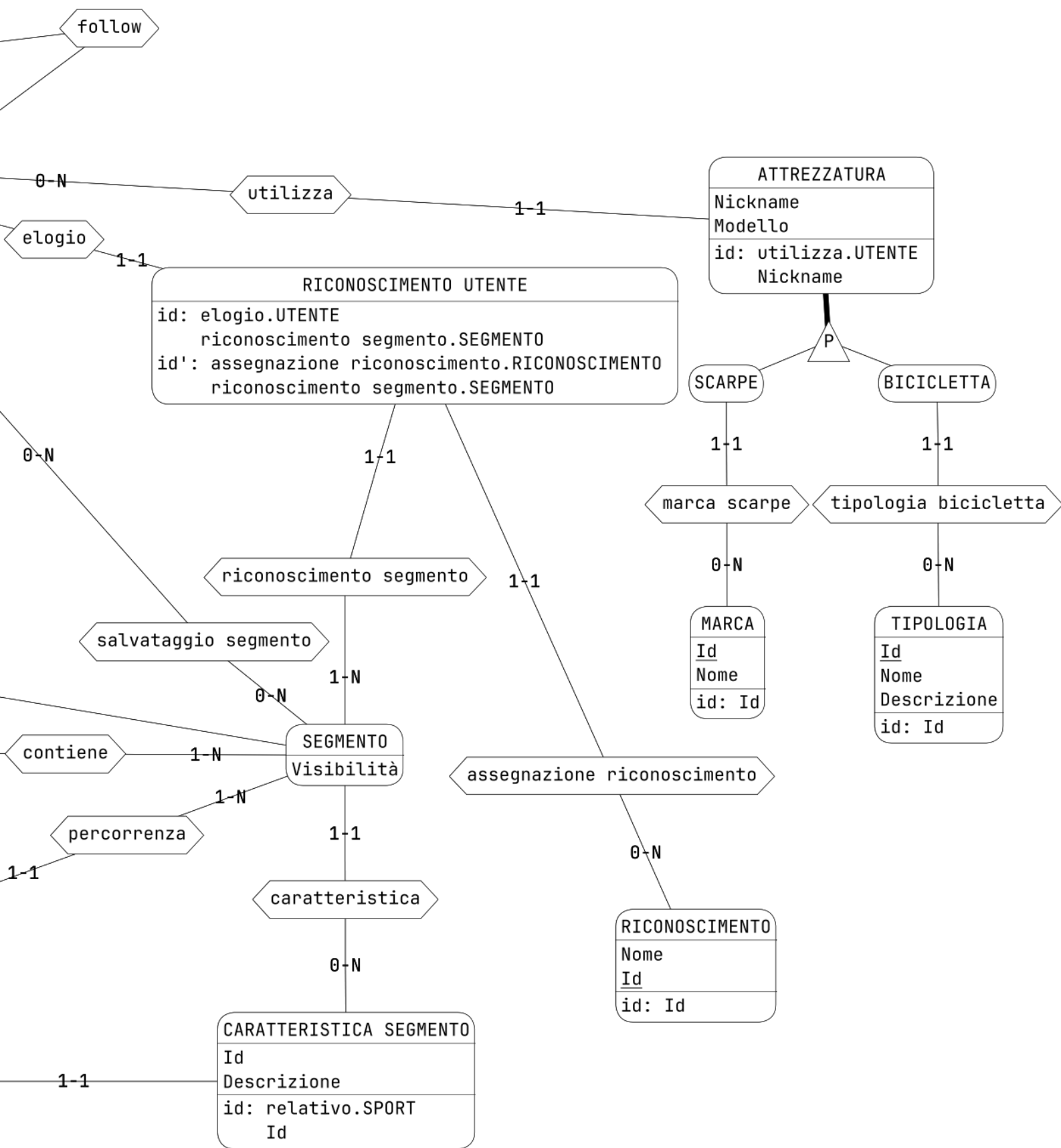
Figura 2.7: *Schema raffinamento riconoscimento.*

I riconoscimenti vengono assegnati agli utenti in base al loro tempo di percorrenza su tale segmento. Su un dato segmento solo un utente può avere un determinato riconoscimento. Su un segmento un utente può avere solo un riconoscimento. Un segmento deve avere almeno un riconoscimento che corrisponde all'utente che ha creato un segmento.

## 2.3 Schema concettuale finale

Pagina successiva.





# Progettazione Logica

## 3.1 Stima dei volumi

La stima dei volumi è stata effettuata considerando i dati che un sistema del genere potrebbe generare da un'utenza ridotta e mediamente attivita in un periodo di un anno e una zona ristretta. Se si considerassero anche utenti più o meno attivi e una zona più ampia i dati varierebbero drasticamente.

Concetto	Costrutto	Volume
Utente	E	20.000
Commento	E	10.400.000
Scrive	A	10.400.000
Riceve	A	10.400.000
Attività	E	2.600.000
Ciclismo	E	1.040.000
Corsa	E	910.000
Nuoto	E	650.000
Svolge	A	2.600.000
Effettuazione	A	2.6000000
Itinerario	A	2.600.000
Percorso	E	2.600.000
Segmento percorso	E	52.000.000
Esecuzione	A	52.000.000
Percorrenza	A	52.000.000
Contiene	A	52.000.000

Concetto	Costrutto	Volume
Segmento	E	15.400.000
Caratteristica	A	15.400.000
Caratteristica segmento	E	20
Relativo	A	20
Sport	E	3
Traccia	E	18.000.000
Attrezzatura	E	26.000
Utilizza	A	26.000
Scarpe	E	14.000
Marca scarpe	A	14.000
Marca	E	200
Bicicletta	E	12.000
Tipologia bicicletta	A	12.000
Tipologia	E	5
Salvataggio percorso	A	25.000
Salvataggio segmento	A	40.000
Utilizza	A	45.000
Riconoscimento utente	E	55.440.000
Elogio	A	55.440.000
Riconoscimento segmento	A	55.440.000
Assegnazione riconoscimento	A	55.440.000
Riconoscimento	E	8
Follow	A	180.000

## 3.2 Descrizione delle operazioni principali e stima della loro frequenza

Alcune tra le operazioni più rilevanti che possono essere svolte sono riassumibili nella seguente tabella insieme alla loro frequenza.

Codice	Operazioe	Frequenza
1	Registrazione di un nuovo utente	40/Giorno
2	Svolgimento di un'attività di ciclismo	3.200/Giorno
3	Scrittura di un commento	24.000/Giorno
4	Inserimento di una scarpa da parte di un utente	18/Giorno
5	Assegnazione di un riconoscimento ad un utente	16.000/Giorno
6	Visualizzazione di tutte le attività dei seguiti da un utente, incluso percorso e sport associati	9.000/Giorno
7	Visualizzazione dell'utente col miglior tempo su un dato segmento	3.500/Giorno
8	Visualizzazione di tutte le attività di nuoto svolte in estate	5/Anno
9	Visualizzazione della distanza totale percorsa in un anno da un utente	2.900/Giorno
10	Visualizzazione delle migliori prestazioni per ogni segmento su un dato utente	2/Mese

### 3.3 Tabelle degli accessi e schemi di navigazione

Si cerca ora di stimare i costi di accesso per le varie operazioni descritte in precedenza. Dove non dovesse risultare sufficiente la sola tabella degli accessi sarà fornito anche uno schema di navigazione. Per il calcolo degli accessi i pesi delle operazioni in scrittura sono considerati doppi rispetto a quelli in lettura.

#### 01 - Registrazione di un nuovo utente

La registrazione di un utente si riassume al semplice inserimento dei dati necessari all'entità utente.

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	S
Totale: $1S \cdot 40 \rightarrow 80$ al giorno			

#### 03 - Scrittura di un commento

Un commento deve essere scritto da un utente e riferito ad un'attività.

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	L
Scrivo	A	1	S
Commento	E	1	S
Ricevo	A	1	S
Totale: $(1L + 3S) \cdot 24.000 \rightarrow 168.000$ al giorno			

#### 04 - Inserimento di una scarpa da parte di un utente

L'inserimento di una scarpa richiede la lettura di tutte le marche presenti per permettere la scelta in fase di aggiunta.



Concetto	Costrutto	Accessi	Tipo
Utente	E	1	L
Utilizza	A	1	S
Attrezzatura	E	1	S
Marca scarpe	A	1	S
Marca	E	200	L
Totale: $(201L + 3S) \cdot 18 \rightarrow 3.726$ al giorno			

## 05 - Assegnazione di un riconoscimento ad un utente

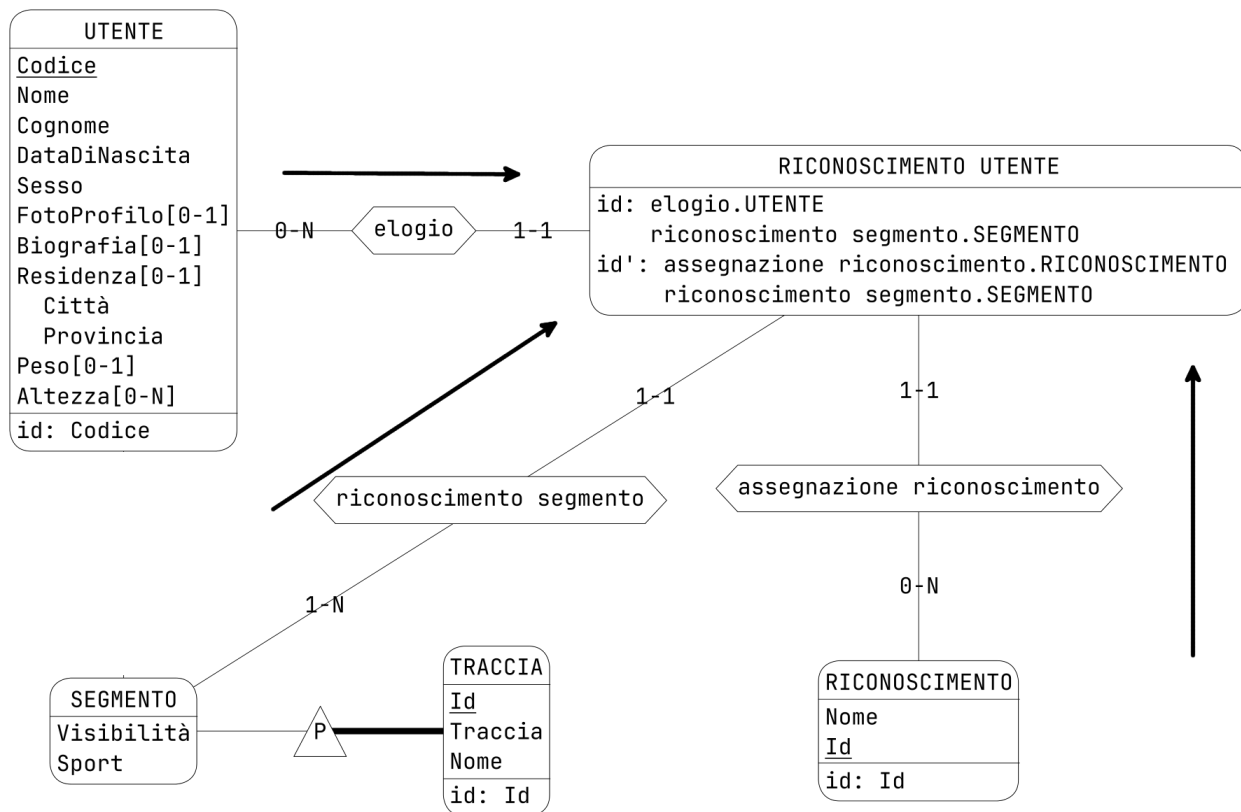


Figura 3.1: Schema di navigazione per l'assegnazione di un riconoscimento.

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	L
Elogio	A	1	S
Segmento	E	1	L
Riconoscimento segmento	A	1	S
Riconoscimento	E	1	L
Assegnazione riconoscimento	A	1	S
Riconoscimento utente	E	1	S
<b>Totale:</b> $(3L + 4S) \cdot 16.000 \rightarrow 176.000$ al giorno			

## 06 - Visualizzazione di tutte le attività dei seguiti da un utente

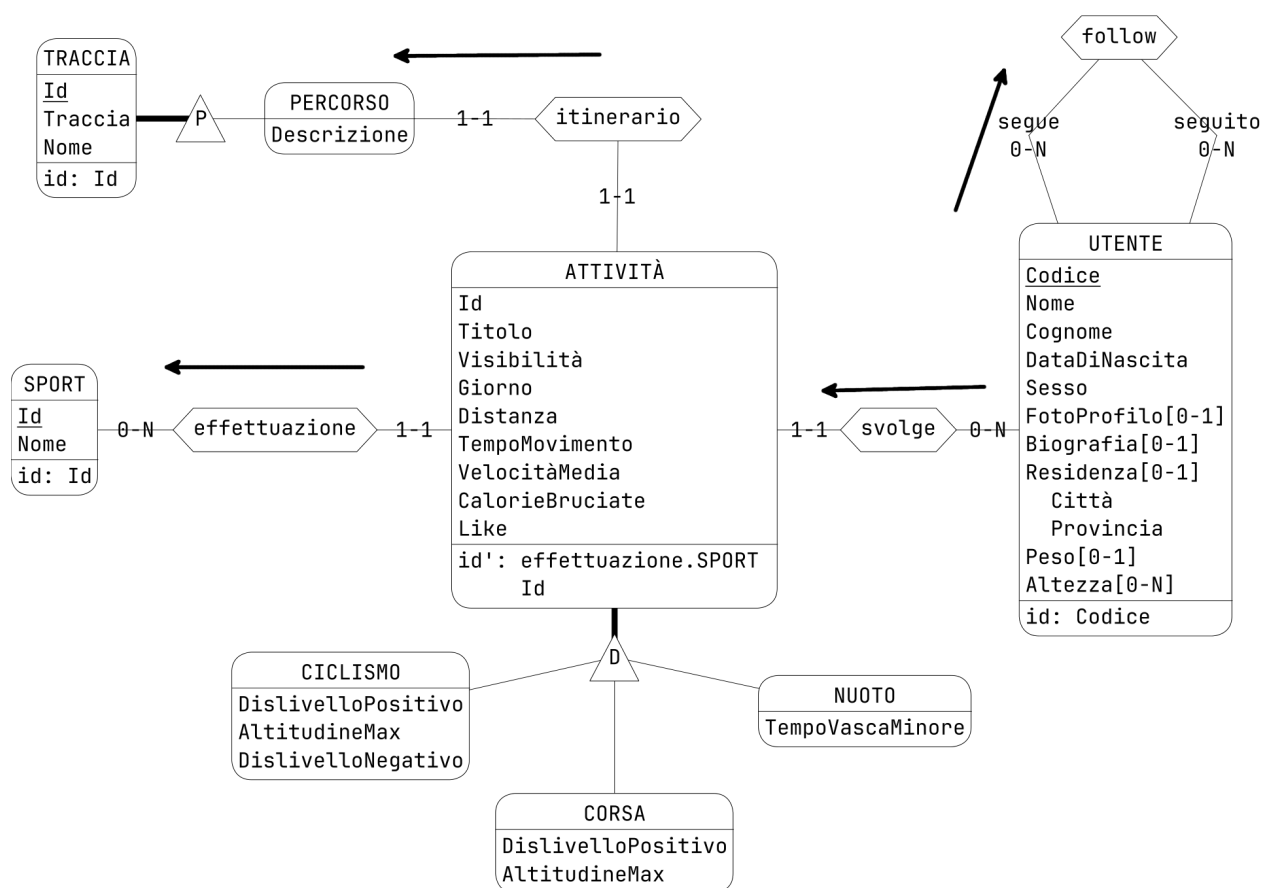


Figura 3.2: Schema di navigazione per le attività dei seguiti.

Supponendo che un utente segua in media 5 altri utenti.

Concetto	Costrutto	Accessi	Tipo
Follow	A	5	L
Utente	E	5	L
Svolge	A	5	L
Attività	E	5	L
Itinerario	A	5	L
Percorso	E	5	L
Effettuazione	A	5	L
Sport	E	5	L
<b>Totale:</b> 40L · 9.000 → 360.000 al giorno			

## 07 - Visualizzazione dell'utente con il miglior tempo su un dato segmento

Ogni segmento viene percorso mediamente poco più di 3 volte.

Concetto	Costrutto	Accessi	Tipo
Percorrenza	A	3	L
Segmento percorso	E	3	L
Esecuzione	A	3	L
Attività	E	3	L
Svolge	A	3	L
Utente	E	3	L
<b>Totale:</b> 6L · 3.500 → 21.000 al giorno			

## 08 - Visualizzazione di tutte le attività di nuoto svolte in estate

Siccome ogni attività ha associata la data di effettuazione le operazioni da svolgere si limitano alla lettura di tutte le attività di nuoto e al controllo di quelle svolte in estate. Se poi si necessita di visualizzare anche i relativi percorsi sarà necessario un processo simile a quello usato per l'operazione 6.

Similmente può essere fatto per l'operazione 9, dove è sufficiente leggere tutte le attività di un singolo utente e fare i calcoli necessari.

Concetto	Costrutto	Accessi	Tipo
Attività	E	650.000	L
Totale: 650.000L · 5 → 3.250.000 all'anno			

## 10 - Visualizzazione delle migliori prestazioni per ogni segmento dato un utente

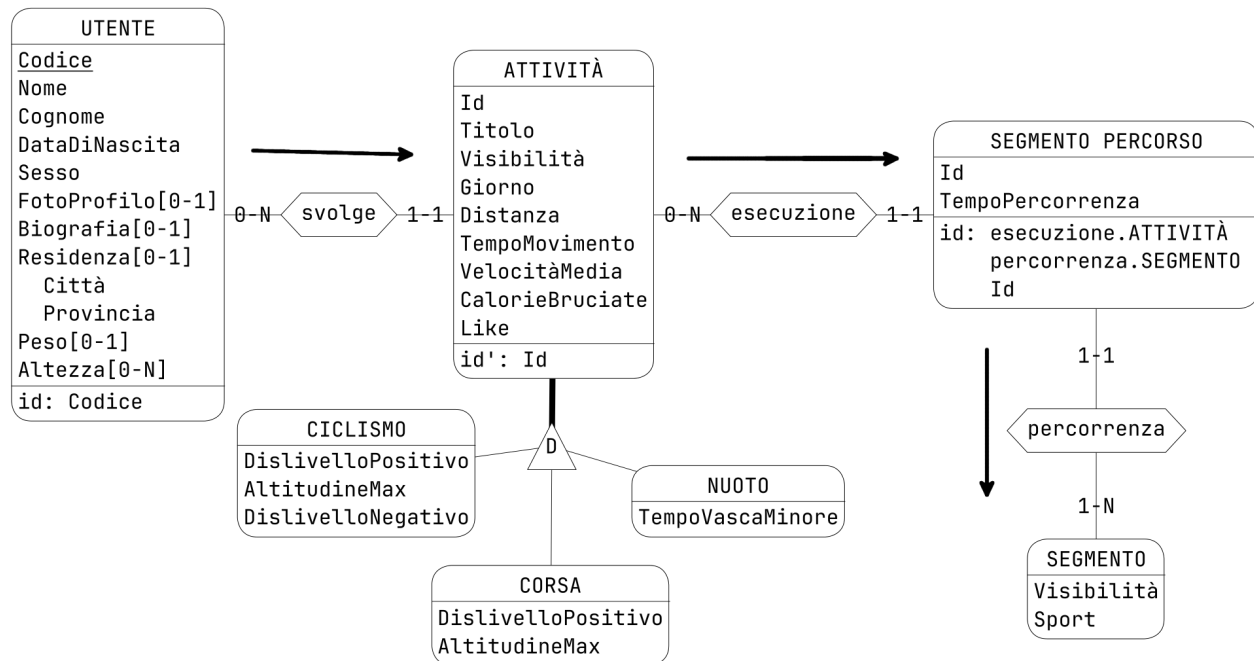


Figura 3.3: Schema di navigazione per le migliori prestazioni.

Ogni utente ha in media 130 attività svolte.

Concetto	Costrutto	Accessi	Tipo
Svolge	A	130	L
Attività	E	130	L
Esecuzione	A	2.600	L
Segmento percorso	E	2.600	L
Percorrenza	A	2.600	L
Segmento	E	2.600	L
Totale: 10.660L · 2 → 21.320 al mese			

### 3.4 Analisi delle ridondanze

#### 02 - Svolgimento di un'attività di ciclismo

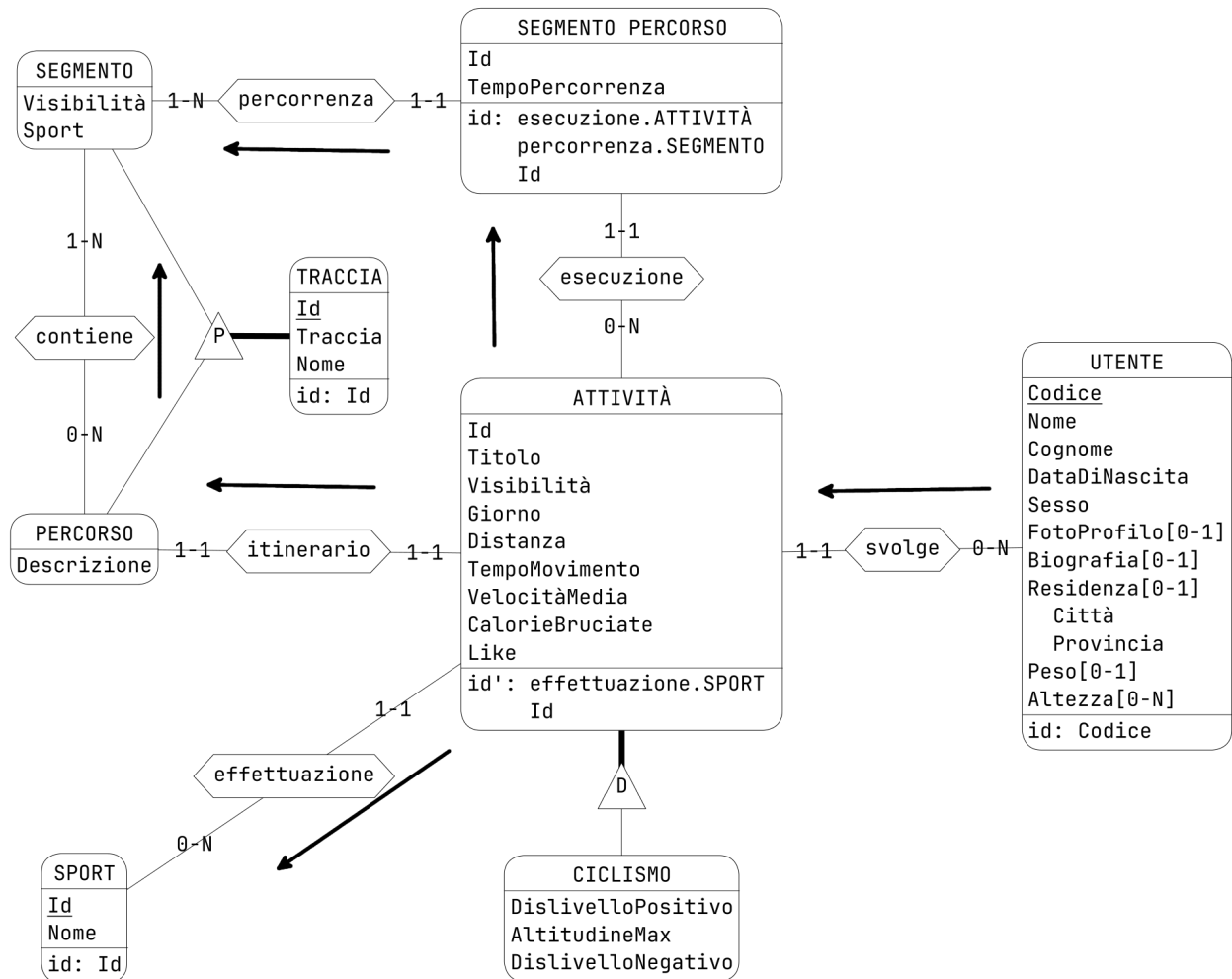


Figura 3.4: Schema di navigazione per lo svolgimento di un'attività.

Ogni percorso contiene circa 20 segmenti.

Concetto	Costrutto	Accessi	Tipo
Utente	E	1	L
Svolge	A	1	S
Attività	E	1	S
Itinerario	A	1	S
Percorso	E	1	S
Effettuazione	A	1	S
Contiene	A	20	S
Esecuzione	A	20	S
Segmento Percorso	E	20	S
Percorrenza	A	20	S
<b>Totale:</b> $(1L + 85S) \cdot 3.200 \rightarrow 547.200$ al giorno			

L'assegnazione ad un percorso dei segmenti corretti richiede l'utilizzo di un algoritmo in grado di analizzarli ed ha come conseguenza che i costi di accesso potrebbero variare in base al numero dei segmenti. Ipotizzando una distribuzione non omogenea nei segmenti tra gli sport, in particolare considerando per il ciclismo 1/2 dei segmenti (7.700.000) e una divisione equa per i restanti:

Concetto	Costrutto	Accessi	Tipo
Segmento	E	7.700.000	L
Caratteristica	A	7.700.000	L
Caratteristica segmento	E	7.700.000	L
<b>Totale:</b> $23.100.000L \cdot 3.200 \rightarrow 7.392 \times 10^{10}$ al giorno			

Per cercare di ridurre il numero degli accessi si è deciso di inserire una ridondanza in segmento. Grazie all'attributo Sport si può evitare di passare per *Caratteristica segmento*.

Concetto	Costrutto	Accessi	Tipo
Segmento	E	7.700.000	L
<b>Totale:</b> $7.700.000L \cdot 3.200 \rightarrow 2.464 \times 10^{10}$ al giorno			

## 3.5 Raffinamento dello schema

### Eliminazione delle gerarchie

Per la gerarchia *Attività* (p, e) si è scelto di adottare il collasso verso l'alto dato che le entità figlie sono accedute in maniera contestuale a quella padre. Risulta inoltre comodo avere tutti i parametri relativi ad un'attività in un'entità unica.

Per la gerarchia *Traccia* (t, e) si è scelto di adottare il collasso verso il basso siccome gli accessi alle entità figlie sono distinti e non si ha necessità che l'identificatore sia univoco. Scelta analoga è stata fatta per la gerarchia *Attrezzatura* (t, e), con differenza nella scelta dell'identificatore.

### Eliminazione degli attributi composti

L'attributo composto *residenza* nell'entità *Utente* è stato scomposto nelle sue componenti, sarà poi necessario controllare a livello applicativo che i valori vengano inseriti in maniera coerente.

### Scelta delle chiavi primarie

Nello schema sono già evidenziate tutte le chiavi primarie per la maggior parte delle entità. Per l'entità *Riconoscimento utente* si sceglie di utilizzare come chiave primaria l'associazione con *Utente* e *Segmento*, lasciando il rimanente vincolo (assegnare un solo riconoscimento per utente) da implementare a livello applicativo.