

Teknologi Basis Data Berorientasi Graf

Ihsan Imaduddin Azhar 13517043
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13517043@std.stei.itb.ac.id

Abstrak—Sudah lebih dari satu dekade manajemen basis data menjadi bahan penelitian. Saat ini ada berbagai jenis sistem manajemen basis data. Pada basis data relasional atau basis data SQL, data disimpan dalam bentuk table. Berbeda dengan basis data relasional, basis data direpresentasikan dengan graf.

Kata kunci—Graf, basis data, *relationship*, representasi data.

I. PENDAHULUAN

Basis data adalah kumpulan informasi yang disimpan didalam computer secara sistematis sehingga pengaksesan dan pembaruan data dapat dilakukan dengan mudah. Basis data sudah menjadi kebutuhan organisasi saat ini.

Basis data tradisional yang saat ini banyak digunakan adalah basis data relasional. Basis data relasional merepresentasikan datanya sebagai baris dan kolom. Baris dan kolom pada basis data relasional dapat diumpamakan sebagai baris dan kolom pada *Excel spreadsheet*. Basis data relasional sudah lama menjadi standar pada industry, terlebih lagi penyedia produk basis data relasional adalah perusahaan teknologi raksasa Oracle dan Microsoft. Basis data relasional baik digunakan untuk beberapa data tapi pemodelan baris dan kolom pada basis data relasional tidak tidak selalu tepat digunakan untuk semua data karena basis data relasional tidak dapat merepresentasikan hubungan antar datum dengan baik. Dengan memisahkan baris dan kolom, basis data relasional sangat sulit diterapkan pada data yang memiliki *interconnection* pada data tersebut. dengan kreativitas memang mungkin data berorientasi-relasional disimpan dalam basis data relasional, tetapi pemodelan tersebut sangat tidak efisien dan lambat dalam pengaksesan data.

Beberapa tahun ke belakang, terlahir tipe basis data baru untuk menyimpan data berorientasi relasional dengan lebih efisien. Nama tipe basis data tersebut adalah basis data graf. Basis data graf (*graph database*) adalah basis data yang menggunakan graf dalam pemodelan data.

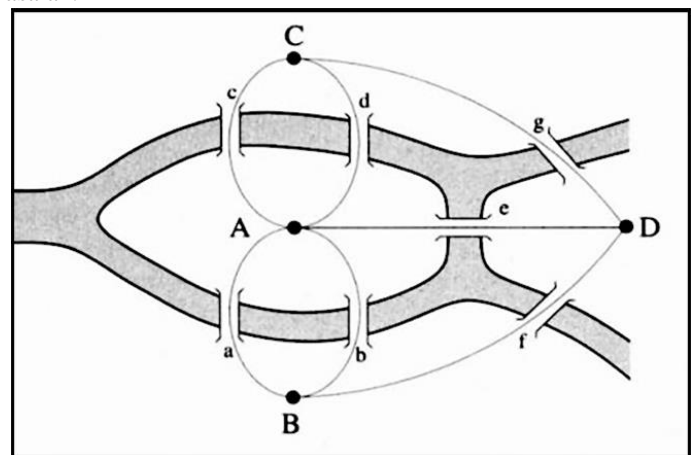
Graf sendiri merupakan cabang ilmu matematika. Teori graf pertama kali muncul pada tahun 1976, yakni ketika Leonard Euler mencoba memecahkan persoalan jembatan Konisberg.

II. LANDASAN TEORI

Untuk memahami penerapan graf dalam basis data dibutuhkan pengetahuan tentang teori graf dan sedikit tentang apa itu basis data.

2.1 Graf

Makalah yang ditulis oleh matematikawan Leonard Euler perihal masalah jembatan Konisberg adalah makalah pertama dalam sejarah teori graf. Semenjak saat itu teori graf berkembang dan diterapkan dalam berbagai pemecahan masalah.



Gambar 2. 1 Persoalan jembatan Konisberg

2.1.1 Definisi Graf

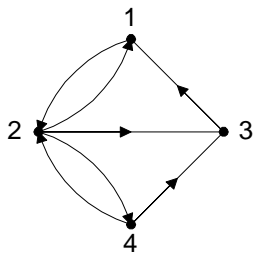
Secara matematis graf didefinisikan sebagai $G=(V,E)$, dimana V adalah himpunan simpul-simpul (vertex atau node) yang tidak kosong dan E adalah himpunan sisi (edge) yang menghubungkan dua simpul.

Menurut definisi sebuah graf tidak boleh memiliki sisi yang kosong, sedangkan himpunan dari sisi suatu graf boleh kosong. Dengan kata lain, sebuah graf minimal memiliki sebuah simpul untuk dikatakan sebagai graf. Graf yang panjang himpunan simpulnya satu dan panjang himpunan sisinya nol disebut graf trivial.

2.1.2 Jenis-jenis Graf

- A. Berdasarkan ada atau tidaknya sisi ganda
 - a. Graf Sederhana
Graf yang tidak mengandung sisi ganda.
 - b. Graf tak-sederhana
Graf yang memiliki sisi ganda.
- B. Berdasarkan orientasi arah pada sisi
 - a. Graf tak-berarah
Graf yang sisinya tidak mempunyai orientasi arah disebut graf tak-berarah.

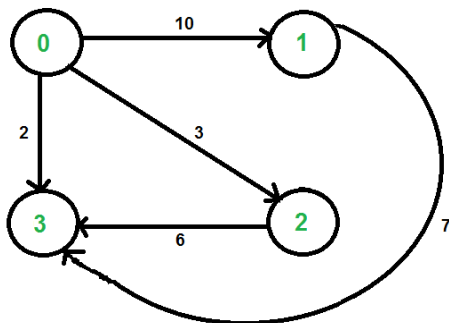
- b. Graf berarah
Graf yang setiap sisinya diberikan orientasi arah.



Gambar 2. 2 graf berarah

C. Berdasarkan ada atau tidaknya bobot pada sisi

- a. Graf berbobot
Graf yang setiap sisinya tidak hanya menghubungkan simpul tapi juga memiliki nilai pada setiap sisinya.
- b. Graf tak-berbobot
Graf yang setiap sisinya hanya menghubungkan simpul dan tidak memiliki nilai pada setiap sisinya.



Gambar 2. 3 Graf berbobot

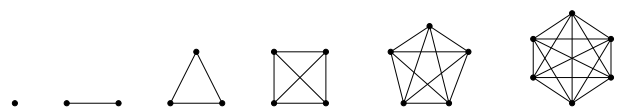
2.1.3 Terminologi pada graf

- Ketetanggaan
Dua buah simpul dikatakan bertetangga bila kedua simpul tersebut dihubungkan langsung oleh sebuah sisi.
- Bersisian
Untuk sembarang sisi e yang menghubungkan v_1 dan v_2 , dapat dikatakan e bersisian dengan v_1 dan v_2 .
- Simpul terpencil
Suatu simpul dikatakan terpencil jika tidak ada sisi yang bersisian dengan simpul tersebut.
- Graf kosong
Suatu graf dikatakan graf kosong jika himpunan sisi dari graf tersebut adalah himpunan kosong.
- Derajat
Derajat suatu simpul adalah jumlah sisi yang bersisian dengan simpul tersebut.
- Lintasan
Lintasan yang sisi yang dilewati dari simpul awal v_0 ke simpul tujuan v di dalam graf G .

- Sirkuit
Lintasan yang berawal dan berakhir di sisi yang sama disebut sirkuit.
- Terhubung
Dua simpul dikatakan terhubung jika terdapat lintasan yang menghubungkan kedua simpul tersebut.
- Upa-graf
Misalkan $G=(V,E)$ adalah sebuah graf. $G_1=(V_1,E_1)$ dikatakan upagraf dari G jika dan hanya jika V_1 adalah himpunan bagian dari V dan E_1 merupakan himpunan bagian dari E .
- Komplemen Upagraf
Komplemen dari sebuah upagraf G_1 terhadap G adalah graf $G_2=(V_2,E_2)$ sedemikian sehingga $E_2= E-E_1$ dan V_2 adalah himpunan simpul yang anggota E_2 bersisian dengannya.
- Upagraf Rentang
Suatu $G_1=(V_1,E_1)$ dapat dikatakan upagraf rentang dari $G=(V,E)$ jika $V_1=V$. Dengan kata lain G_1 mengandung semua simpul yang ada pada G .
- Cut-set
Cutset dari graf terhubung $G=(V,E)$ adalah himpunan bagian dari E yang jika anggotanya dihilangkan membuat graf G tidak lagi terhubung.
- Graf Berbobot
Graf berbobot adalah graf yang setiap sisinya diberi nilai.

2.1.4 Beberapa graf khusus

1. Graf lengkap
Graf lengkap adalah graf yang mempunyai sisi ke semua simpul lainnya.



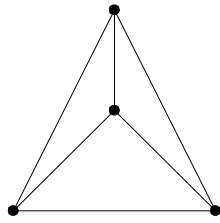
Gambar 2.4 Graf lengkap

2. Graf lingkaran
Graf lingkaran sederhana adalah graf yang setiap simpulnya hanya memiliki dua simpul (berderajat dua).



Gambar 2. 5 Graf Lingkaran

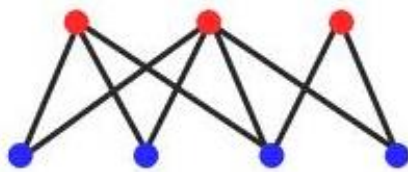
3. Graf teratur
Graf teratur adalah graf yang setiap simpulnya memiliki derajat sama.



Gambar 2. 6 Graf teratur

4. Graf bipartite

Graf bipartite adalah graf yang setiap simpulnya dapat dikelompokkan menjadi dua bagian V_1 dan V_2 , sedemikian sehingga setiap sisi pada G menghubungkan sebuah simpul pada V_1 ke V_2 .



Gambar 2. 7 Graf bipartite

2.2 Basis Data

2.2.1 Definisi Basis Data

Basis data adalah koleksi informasi yang terorganisir sehingga dapat diakses, di-manaje, dan diperbarui.

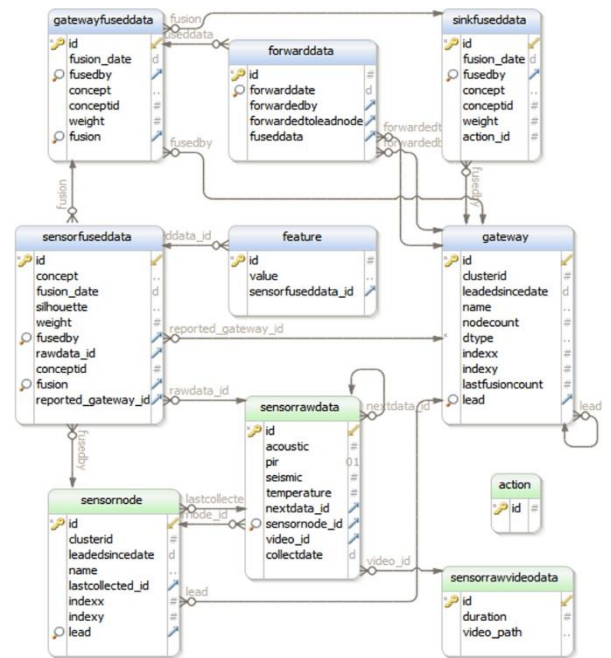
2.2.2 Pentingnya basis data

Basis data dapat menghasilkan informasi yang sangat penting untuk pemilik data tersebut, baik perusahaan maupun individu, data tersebut harus dikelola dengan baik, terpadu dan juga aman. Pengelolaan data yang baik adalah ketika data yang dikelola tersebut mudah diakses oleh pemilik datanya. Sedangkan terpadu adalah ketika pengelolaan data tersebut berkesinambungan. Sedangkan pengelolaan data yang aman adalah data yang dapat melindungi yang dikelola terlindungi dari pihak lain,

2.2.2 Beberapa Jenis Basis Data

A. Basis Data Relasional

Basis data relasional ditemukan oleh E.F.Codd pada tahun 1970. Basis data relasional adalah basis data yang model datanya didasarkan oleh model data relasional. Hampir semua basis data relasional menggunakan SQL (*Structured Query Language*) sebagai Bahasa untuk *query*-nya.



Gambar 2. 8 Skema basis data relasional

B. Basis Data Terdistribusi

Basis data terdistribusi adalah basis data yang beberapa porsi datanya tersimpan pada berbagai lokasi fisik. Basis data terdistribusi dapat dibedakan menjadi dua yaitu basis data homogeny dan heterogen. Semua lokasi fisik pada basis data terdistribusi homogen memiliki perangkat keras dan system operasi yang sama. Basis data heterogen sebaliknya.

C. Basis Data Graf

Basis data berorientasi graf adalah basis data yang menggunakan graf berarah dalam pemodelan data. Pada basis data graf, data dimodelkan dalam bentuk graf. Graf ini memiliki sesuatu yang disebut *nodes* (Simpul) yang dihubungkan oleh *relationship* (sisi). Karena representasi visualnya berupa graf berarah, basis data graf baik digunakan untuk menganalisis pola dari suatu data yang besar.

D. Basis data berorientasi dokumen

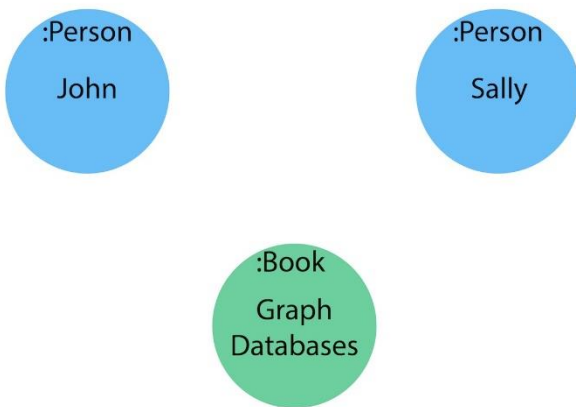
Basis data berorientasi dokumen adalah basis data yang datanya disimpan dalam bentuk dokumen yang biasanya dalam format XML, JSON maupun BSON. Sifat dokumen pada basis data berorientasi dokumen semi terstruktur dan tidak disimpan di table seperti basis data relasional. Basis data berorientasi dokumen tidak memiliki schema sehingga dokumen yang berbeda beda dapat disimpan pada basis data tersebut. basis data berorientasi dokumen sangat berguna menyimpan yang strukturnya dinamis atau mudah berubah. Basis data berorientasi dokumen banyak digunakan perusahaan skala besar.

III. BASIS DATA GRAF

Basis data berorientasi graf adalah basis data yang menggunakan graf berarah dalam pemodelan data. Pada basis data graf, data dimodelkan dalam bentuk graf. Graf ini memiliki sesuatu yang disebut *nodes* (Simpul) yang dihubungkan oleh *relationship* (sisi).

3.1 Nodes

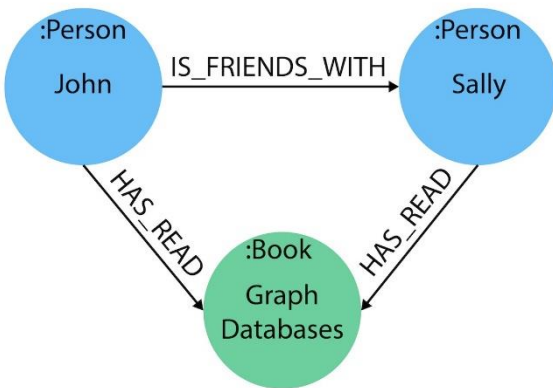
Setiap *node* pada basis data graph dapat memiliki atribut. Atribut tersebut disebut dengan properti. Setiap *node* pada basis data graf memiliki label yang menandakan apa peran *node* tersebut dalam data tersebut. Misalkan data yang dimodelkan dengan basis data graf adalah data hasil survey, label yang mungkin digunakan adalah “partisipasi” dan “jawaban”.



Gambar 3. 1 Contoh Nodes

3.2 Relationship

Relationship adalah sisi berarah yang menghubungkan *nodes*. Seperti *nodes*, hampir semua *relationship* pada basis data graf memiliki attribute, yang disebut property. Properti tersebut menentukan apa hubungan antara dua *nodes* yang dihubungkan oleh *relationship* tersebut. Dalam dunia perkuliahan, salah satu *relationship* yang mungkin adalah “teman dari”.



Gambar 3.2 Contoh relationship yang menghubungkan nodes

3.3 Query

Dalam teknologi basis data graf, ada beberapa bahasa query yang populer digunakan.

3.3.1 Cypher

Cypher adalah *declarative query language* yang dipopulerkan oleh Perusahaan Neo4j. Bahasa query ini populer karena kemiripannya dengan SQL.

```
1 MATCH (n:Person)-[:FRIEND]-(f)
2 WITH count(f) as c, n
3 MATCH (n)-[:FRIEND]-(f)-[:FRIEND]-(fof)
4 RETURN n, c, fof
```

Gambar 3. 4 Query Cypher

3.3.2 GraphQL

GraphQL adalah Bahasa query yang dipopulerkan oleh facebook. GraphQL adalah Bahasa query untuk API yang tidak spesifik hanya untuk basis data graf. Pada Bahasa query GraphQL, pengguna member struktur data yang mereka butuhkan dan akan mengembalikan apa nilai struktur data yang telah diberikan pengguna.

```
1 GraphQL Query
2 {
3   hero {
4     name
5     mass
6   }
7 }
8 Returns
9 {
10  "hero": {
11    "name": "Luke Skywalker",
12    "mass": 77
13  }
14 }
```

Gambar 3. 4 Query GraphQL

Dengan menambahkan atribut, pengguna mendapatkan hasil yang berbeda.

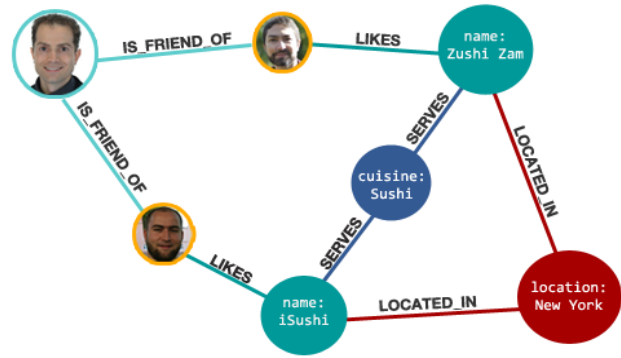
```
1 {
2   hero {
3     name
4     height
5     mass
6   }
7 }
8 Returns
9 {
10  "hero": {
11    "name": "Luke Skywalker",
12    "height": 1.72
13    "mass": 77
14  }
15 }
```

Gambar 3. 5 Query GraphQL setelah atribut ditambahkan

3.4 Kasus penggunaan basis data graph

Basis data graf baik digunakan ketika data yang ingin direpresentasikan memiliki hubungan antar datum yang terhubung kuat atau *highly connected*. Sangat cocok untuk merepresentasikan data seperti *social network*.

Basis data graf juga sangat baik digunakan untuk data yang *relationship* antar *nodes*-nya sama pentingnya dengan *nodes* itu sendiri seperti data perilaku konsumen. Oleh karena itu basis data graf banyak ditemukan pada mesin rekomendasi.



Gambar 3. 6 Contoh graf rekomendasi

IV. KELEBIHAN BASIS DATA GRAF

A. Performa

Karena basis data graf pada dasarnya merupakan struktur data indeks, untuk pencarian query tidak dibutuhkan untuk memuat data yang tidak berhubungan. Dan hal tersebut cocok untuk *big data analytical query*.

B. Fleksibel

Basis data graf fleksibel karena properti, *node*, dan *edge* baru dapat ditambahkan dan dihilangkan kapanpun. Tidak perlu membuat table baru untuk menambahkan properti dari basis data graf.

C. Familiar

Node adalah sesuatu yang programmer berorientasi objek sangat familiar. *Node* dapat merepresentasikan orang, bisnis, akun atau benda apapun yang kita mau.

V. APLIKASI BASIS DATA GRAF

A. Mesin rekomendasi

Salah satu kegunaan basis data graf yang populer adalah mesin rekomendasi.

Mesin rekomendasi digunakan dalam berbagai bidang industry untuk meningkatkan pengalaman belanja pelanggan dengan menebak apa yang pelanggan tersebut inginkan selanjutnya.

Mesin rekomendasi tradisional bekerja secara offline dengan memproses sejarah transaksi tiap pelanggan dengan algoritma yang kompleks dan menghasilkan rekomendasi mingguan atau rekomendasi bulanan.

Pada saat pertama kali diperkenalkan, rekomendasi mingguan adalah keuntungan besar bagi yang mengadopsi teknologi tersebut. Namun karena teknologi tersebut saat ini telah tersebar luas, tidak cukup lagi hanya bergantung pada mesin rekomendasi tradisional.

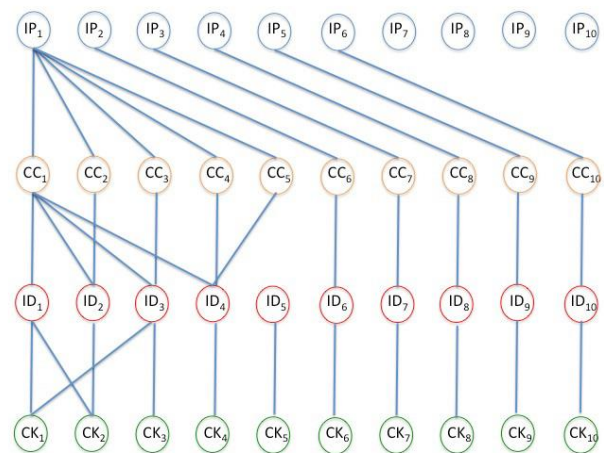
Oleh karena itu, basis data graf cocok digunakan sebagai mesin rekomendasi yang lebih canggih karena pada basis data graf kita dapat melakukan *relationship-based query* atau query berdasarkan *relationship*. Dengan hasil dari *relationship-based query*, kita dapat menganalisis pola perilaku belanja pelanggan dengan mudah.

B. Deteksi Penipuan

Semenjak kehidupan menjadi semakin terdigitalisasi dan system pembayaran online diterapkan oleh banyak pelaku usaha, jumlah transaksi online semakin pun bertumbuh. Untuk memanfaatkan hal tersebut, para penipu beradaptasi dengan cepat, oleh karena itu dibutuhkan harus ada cara untuk mengatasi permasalahan penipuan tersebut.

Kegiatan penipuan seperti ini biasanya melibatkan identitas palsu dalam jumlah banyak,

Misalkan dalam praktik transaksi online hal yang dapat dimonitor adalah ID, alamat IP, Cookie dan kartu credit pengguna tersebut. Jika data tersebut kita simpan dalam *nodes*, pada umumnya *relationship* antar *nodes* tersebut adalah hubungan *one-to-one*. Dari pola tersebut, jika *relationship* antara terhitung banyak secara signifikan, dapat ditarik kesimpulan kemungkinan besar pengguna tersebut adalah penipu.



Gambar 5. 1 graf transaksi online

Dari gambar 5.1 dapat ditarik kesimpulan pengguna dengan alamat IP₁ kemungkinan besar adalah penipu karena pengguna menggunakan lebih dari satu kartu credit dan terdapat banyak *interconnected relationship*.

V. KESIMPULAN

Teori graf merupakan teori yang sudah sangat lama, tapi hingga saat ini teori graf masih bisa dikembangkan untuk berbagai masalah pada dunia teknologi maupun sehari-hari.

Penerapan teori graf dalam basis data graf menghadirkan solusi untuk masalah masalah yang ada pada pemodelan basis data selama ini dan dapat diterapkan berbagai bidang. Walaupun basis data graf merupakan teknologi baru yang masih memiliki banyak kekurangan, basis data graf memiliki potensi yang besar terutama di zaman informasi *big data* seperti saat ini.

VII. ACKNOWLEDGMENT

Penulis berterimakasih kepada Dr. Rinaldi Munir sebagai dosen pengampu IF2120, mata kuliah Matematika Diskrit, untuk memberikan kesempatan kepada saya untuk menulis makalah ini. Penulis juga berterimakasih kepada orang tua penulis karena telah memberikan dukungan moral kepada saya selama masa penulisan, penulis juga berterimakasih kepada penulis dari referensi dibawah untuk kontribusi mereka dalam ilmu computer.

REFERENCES

- [1] Munir, Rinaldi. Matematika Diskrit, Bandung: Informatika, 2012, edisi kelima..
- [2] <https://math.stackexchange.com/questions/1173328/eulers-solution-of-seven-bridges-of-k%C3%B6nigsberg-in-layman-terms>
- [3] <https://www.geeksforgeeks.org/graph-implementation-using-stl-for-competitive-programming-set-2-weighted-graph/>
- [4] https://www.researchgate.net/figure/Relational-Database-Schema_fig7_318761128
- [5] <https://developer.ibm.com/dwblog/2017/overview-graph-database-query-languages/>
- [6] <https://www.ontotext.com/5-key-drivers-graph-databases-gaining-popularity/>
- [7] <https://neo4j.com/developer/guide-build-a-recommendation-engine/>
- [8] <https://database.guide/whats-the-difference-between-graph-database-and-relational-database/>
- [9] <https://neo4j.com/developer/guide-data-modeling/>

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 10 Desember 2018



Ihsan Imaduddin Azhae