بسم الله
الرحمن الرحيم

دانشگاه صنعتی اصفهان
دانشکده برق و کامپیوتر

مدل داده شی‌گرای دوزمانی

پایان نامه کارشناسی ارشد مهندسی کامپیوتر- نرم افزار

هادی الیاسی

استاد راهنما
دکتر محمدعلی منتظری

۱۳۹۰

دانشگاه صنعتی اصفهان
دانشکده برق و کامپیوتر

پایان نامه کارشناسی ارشد رشتهٔ مهندسی کامپیوتر هادی الیاسی

تحت عنوان

مدل داده شی‌گرای دوزمانی

در تاریخ ۱۳۹۰/۰۲/۱۹ توسط کمیته تخصصی زیر مورد بررسی و تصویب نهایی قرار گرفت.

۱- استاد راهنمای پایان نامه — دکتر محمدعلی منتظری

۲- استاد داور — دکتر علی فانیان

سرپرست تحصیلات تکمیلی دانشکده — دکتر سید محمود مدرس هاشمی

سپاس‌گزاری

پایان این دوره را مدیون استادانم، دوستانم، خانواده ام هستم.
مدیون و سپاس‌گزار استاد بزرگوارم جناب آقای دکتر منتظری که با راهنمایی دانشورانه شان پایان نامه به اتمام رسید. از ایشان برای تمام زحمات متحمل شده و آموزشها و پیگیری‌ها تشکر می نمایم. از جناب آقای دکتر سرایی که در حل مسائل پوشش و مسیریابی راهنمای من بود تشکر می نمایم. از کارمندان و مسئولین دانشکده مهندسی برق و کامپیوتر برای احساس مسئولیت و یاری بنده در انجام روال‌های اداری تشکر می نمایم.
از دوستانم که هر یک به نوعی من را در طول راه کمک و یاری نمودند، بسیار سپاسگزارم.
سپاسگزارِ همسر مهربانم می‌باشم که همیشه یاور و امید دهنده به من بوده است. از خانواده ام مادر، پدر و برادرم و خواهرم که همیشه پشتیبان و مایه دلگرمیم بوده اند تشکر و قدردانی می نمایم.
در پایان از تمامی کسانی که به هر نحوی مرا یاری نموده اند تشکر و قدردانی می نمایم و امیدوارم که در تمامی مراحل زندگی پیروز و سربلند و تندرست باشند.

هادی الیاسی

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و
نوآوریهای ناشی از تحقیق موضوع این پایان نامه متعلق به
دانشگاه صنعتی اصفهان است.

فهرست مطالب

عنوان ..... **صفحه**
چکیده ۱
1 فصل اول مقدمه ۲
۱-۱- کلیات ۲
۱-۲- شرح مسئله ۳
۱-۳- روش حل مسئله ۳
۱-۴- روند ارائه مطالب ۴
۲ فصل دوم ادبیات تحقیق ۶
۲-۱- مقدمه ۶
۲-۲- داده، اطلاع و موجودیت ۷
۲-۲-۱- داده ۷
۲-۲-۲- اطلاع ۷
۲-۲-۳- موجودیت ۷
۲-۳- انواع دادگان ها ۸
۲-۳-۱- مدل های پایگاه داده ۸
۲-۳-۲- سیستم مدیریت پایگاه دادگان ۱۲
۲-۴- پایگاه داده مبتنی بر زمان ۱۴
۲-۴-۱- ابعاد و واحدهای زمان در پایگاه داده مبتنی بر زمان ۱۴
۲-۴-۲- مدل دادهای زمانی ۱۶
۲-۵- پایگاه داده شی ۲۰
۲-۵-۱- مزایا و معایب پایگاه داد شی ۲۰
۲-۵-۲- زبان های برنامه نویسی شی گرا و ارتباط آن با پایگاه داده شی گرا ۲۲
۲-۵-۳- ویژگی های پایگاه داده ای شی گرا ۲۳
۲-۵-۴- مدل شی پیشنهادی ODMG ۲۴
۲-۶- جمع‌بندی ۲۷
3 فصل سوم مدل شی‌گرا و رابطه‌ای مبتنی بر زمان ۲۹
۳-۱- مقدمه ۲۹
۳-۲- معرفی مدل TOODM ۳۰

۳-۲-۱- فرضیات ۳۰
۳-۲-۲- ساختارها ۳۱
۳-۲-۳- محدودیت ها و قیود ۳۵
۳-۲-۴- عملگرها ۳۶
۳-۳- مدل داده ای UTDM ۳۷
۳-۳-۱- روابط مفهومی دوزمانی ۳۷
۳-۳-۲- شمای نمایش مهرزمانی تاپلی Snodgrass ۴۳
۳-۳-۳- شمای نمایش انبارهای Jenson ۴۷
۳-۳-۴- شمای نمایش مهرزمانی مقدار صفت Gadia ۵۰
۳-۳-۵- شمای نمایش مهرزمانی مقدار صفت Mckenzie ۵۴
۳-۳-۶- شمای نمایش مهرزمانی تاپلی Ben-Zui ۵۹
۳-۴- جمع‌بندی ۶۳
۴ فصل چهارم BTOODM مدل پیشنهادی ۶۴
۴-۱- مقدمه ۶۴
۴-۲- مدل پیشنهادی ۶۵
۴-۲-۱-GenerateVersion ۶۷
۴-۲-۲-UpdateObject ۶۸
۴-۲-۳-GetObjectByTime ۷۴
۴-۲-۴-ObjectSubClass ۷۶
۴-۲-۵-ConvertType ۷۷
۴-۲-۶-GetAttributesValue ۷۸
۴-۲-۷-Rollback ۸۰
۴-۲-۸-GetObjectRelation ۸۱
۴-۳- مقایسه مدل پیشنهادی با سایر مدلهای موجود ۸۴
۴-۳-۱- مدل TOODM و مقایسه با مدل پیشنهادی ۸۴
۴-۳-۲- مدل UTDM و مقایسه با مدل پیشنهادی ۸۵
۴-۳-۳- مدل T_Chimera و چند مدل دیگر و مقایسه با مدل پیشنهادی ۸۵
۴-۴- جمع‌بندی ۸۸
5 فصل پنجم نتیجه‌گیری و پیشنهادات ۸۹

۵-۱- تحقیقات انجام شده ۸۹
۵-۲- پیشنهادات ۹۱
مراجع ۹۳

چکیده

با گسترش روز افزون داده‌ها و اهمیت آنها، محققان مدل‌های داده ای بسیاری در پایگاه‌های داده‌ای ارایه کرده‌اند که هر یک از این مدل‌ها بخشی از مفاهیم را پوشش داده است. با توجه به توانمندی‌های بیشمار شی‌گرایی، در زمینه پایگاه دادگان شی‌گرا مدل‌های بسیار معرفی شده‌اند تا فضاهای خالی میان ویژگی‌های یک شی و مدلی که آن را در پایگاه داده‌ای ذخیره می‌کردند را تقلیل بخشند. ارائه چنین مدل-هایی موجب گردید تا دیگر نیازی به استفاده از نگاشت شی رابطه‌ای(ORM) نباشد. از دیگر سو در طی چند دهه اخیر اهمیت زمان در داده‌ها ملموس‌تر از پیش شده است و اکنون نقش آن در مدل‌های داده‌ای پررنگ‌تر گشته است. مدل شی‌گرایی به دلیل توانایی در انعطاف پذیری و کپسوله سازی از مدل‌هایی محسوب می شود که در چند دهه اخیر بسیار مورد توجه طراحان مدل‌های داده ای قرار گرفته است. از اینرو در این تحقیق مدلی شی‌گرا ارائه گردیده که ابعاد زمان را در خود دارا می‌باشد. البته مدل‌های متعددی از قبیل OODAPLEX، Chimera، TIGUKAT و غیره، توسط محققان مطرح گردیده است. مدل‌های از پیش ارائه شده، در حوزه پایگاه‌های داده‌ای زمانی با محوریت دو بعد از زمان که عبارتند از زمان اعتبار و زمان تراکنش مطرح گشته‌اند. بسیاری از مدل‌ها زمان اعتبار را مد نظر قرار داده‌اند و زمان تراکنش را پشتیبانی نمی‌کند. اهمیت این مسأله تا حدی است که پایگاه‌های داده‌ای که چنین قابلیتی را دربر دارد، Rollback Database نامیده‌اند. برخی نیز تنها از زمان اعتبار پشتیبانی می‌کنند و برخی دیگر نیز هر دو بعد را پوشش می دهند. اما با این وجود مدلی که مفاهیم جامع یک مدل مبتنی بر زمان را بیان کند و بتواند از قابلیت پشتیبانی شی‌های پیچیده بهره ببرد، هنوز ارائه نشده است. از اینرو ما مدلی را مطرح کرده‌ایم که توانمندی‌های بیشتری از جنبه‌های زمانی را در نظر می‌گیرد و قابلیت پشتیبانی از هر دو نوع زمان در مدل‌های داده‌ای- نقطه زمان و بازه زمانی- را دارد.

کلمات کلیدی: پایگاه داده شی‌گرا، پایگاه داده دوزمانی، پایگاه داده زمانمهر، نقطه زمان، بازه زمان، زمان اعتبار، زمان تراکنش

فصل اول
مقدمه

۱-۱- کلیات

مدل کردن داده[1] روشی برای توصیف داده‌ها و عملیات روی آن‌ها در سطوح مختلف معماری پایگاه داده است. یک مدل داده، پدیده‌های دنیای واقعی را توصیف می‌کند و ساختار داده مورد نیاز برای حل مسئله خاص و ارتباط درونی بین ساختارها را تعریف می‌کند[1][2]. تا کنون مدل‌های داده‌ای متعددی ارائه شده‌اند که هر یک دارای نقاط قوت وضعفی می‌باشند[3]، اما در این بین مدل داده شی‌گرا از آن دسته مدل‌هایی محسوب می‌شود که به نسبت دیگر مدل‌ها توانایی بیشتری در مدل کردن اشیای موجود در جهان واقعی را دارا است. در ابتدا چالش‌های بسیاری در این زمینه وجود داشت، اما پس از مدت کوتاهی کنسرسیومی با نام ODMG تشکیل گردید که با تعریف یک استاندارد این مشکلات را تا حد زیادی برطرف نمود.

در یک مدل شی‌گرا هر شی مشخصه یکتای خودش را دارد وحتی زمانیکه از سیستم حذف گردد نیز شناسه آن به شی جدید تخصیص داده نمی‌شود. در این مدل هر شی می‌تواند مالک چند شی دیگر باشد و همین ویژگی آن است که این مدل را بر دیگر مدل‌ها همانند رقیب سرسختش، مدل رابطه‌ای برتری می‌دهد.مفهوم مالکیت بدین معنی است که به عنوان مثال شیئی با نام رضا می‌تواند مالک شی آدرس، خودرو و شغل و ... نیز باشد. مفاهیم دیگر موجود

[1] Data Modeling

در مدل شی‌گرا از قبیل وراثت و کپسوله‌سازی و غیره از مواردی هستند که نکته‌ای قابل اتکا در انتخاب این مدل برای ما می‌باشد. اما در مقابل پیچیدگی استفاده از آن و به تبع آن پایین بودن قابلیت اطمینان به دلیل موارد استفاده اندک برای شرکت‌ها موجب می‌شود تا گروه اندکی به سوی این مدل گام بردارند[4][5][6].

با گذشت زمان و افزایش بی‌رویه حجم داده‌ها، نیاز به استفاده از داده‌هایی که از پیش وجود داشتند محسوس‌تر شد و این مسأله‌ای بود که محققان را برآن داشت تا توجه ویژه‌ای به نقش زمان در پایگاه‌های داده‌ای داشته باشند. از اینرو مدل داده‌ای با نام مدل داده زمانی[1] عرضه گشت تا دیگر زمان به‌صورت یک بعد در سیستم وجود داشته باشد و دیگر نیازی به تعریف آن از سوی کاربر نباشد. بنابراین محققان بر آن شدند تا مدل داده‌ای شی-گرای زمانی را ارائه نمایند تا هم قابلیتهای فراوان شی را در بر داشته باشد و هم زمان را به عنوان بعدی جدایی ناپذیر از داده‌ها در خود داشته باشد[7][8][9][10].

۱-۲- شرح مسئله

هدف از انجام این پروژه ارائه یک مدل کارآمد با استفاده از دو مدل شی‌گرا و مدل زمانی می‌باشد. در چند دهه اخیر مدل‌های متعدد زمانی مبتنی بر مفاهیم و اصول شی‌گرایی ارائه شده است که هر یک روی بخشی از قابلیت-ها تمرکز کرده‌اند[11] [12] [13] [14] [15]. از آنجاکه زمان در سیستم‌های پایگاه داده‌ای دارای دو بعد زمان اعتبار و زمان تراکنش می‌باشد، برخی از مدل‌ها زمان اعتبار [11] [12] [13] و برخی دیگر تنها بعد زمان تراکنش و برخی هر دو بعد[2] [14][15] را پشتیبانی می‌نمایند.

از دیگر سو به لحاظ فاصله و بازه می‌توان زمان را در یک مدل به دو دسته نقطه زمان[3] و بازه زمان[4] تقسیم می‌گردد و این خود موجب شده است تا برخی از مدل‌ها تنها از نوع اول و برخی دیگر تنهااز نوع دوم استفاده کنند. البته مدل‌هایی هم ارائه شده‌اند که تمامی حالت‌ها را پوشانیده باشند، اما از ویژگی‌های مفید شی‌گرایی برای مدیریت اشیا پیچیده به درستی بهره نبرده است.

۱-۳- روش حل مسئله

پایان‌نامه حاضر پس از مطالعه موارد مشابه و مفاهیم بنیادی در رابطه با مسئله تعریف شده روتین‌هایی را ارائه کرده است تا با استفاده از آن‌ها بتوان هر دونوع فواصل زمانی و هر دو بعد از زمان را پشتیبانی نمود. علاوه بر آن این متدها به گونه‌ای طراحی شده‌اند تا از قابلیت شی‌گرایی در حد مطلوب استفاده شود. استفاده از انواع داده‌ای

---

[1] Temporal Data Model
[2] Bi-Temporal
[3] Time Point
[4] Time Interval

متعدد(Bag, Collection, Dictionary, List, Set) موجب شده است تا مدیریت شی پیچیده در سیستم به نحو موثری انجام گیرد. ارائه هشت روتین که امکان بهنگام سازی مقادیر صفت‌ها بسته به دو بعد از زمان، ایجاد یک نسخه جدید از یک شی در یک لحظه خاص، امکان مشاهده لیستی از اشیا در هر دو بعد از زمان، امکان مشاهده کلاس‌های موروثی از یک کلاس، تبدیل نوع صفات مربوط به یک شی، امکان مشاهده مقادیر صفات از یک شی خاص، امکان Rollback کردن به زمانی معتبر، مشاهده وضعیت دو شی نسبت به یکدیگر با استفاده از مدل معرفی شده Allen [16] توسط مدل مطرح شده مهیا شده است تا مدلی به نسبت کامل ارائه گردد.

۱-۴- روند ارائه مطالب

در این پایان نامه، در ابتدا و در فصل دوم مروری بر مفاهیم پایه‌ای پایگاه‌های داده مانند داده و اطلاع و موجودیت خواهیم داشت. تفاوت‌ها و جایگاه هریک در سیستم‌های داده‌ای را بررسی می‌کنیم. سپس در ادامه مدل داده و ویژگی‌های آن را بر می‌شمریم. انواع مختلف مدل‌های داده‌ای اعم از مدل تخت و شبکه‌ای وسلسله مراتبی و شی گرا را مورد بررسی قرار می‌دهیم. سپس از بین این مدل‌ها دو مدل مرتبط با حوزه کاری این پایان‌نامه یعنی مدل شی‌گرا و مدل مبتنی بر زمان را بیشتر و کامل‌تر معرفی می‌نماییم. در معرفی مدل زمانی مفاهیم کلی از قبیل ابعاد زمان در پایگاه داده که مشتمل بر دو نوع زمان اعتبار و زمان تراکنش می‌باشد، فواصل زمانی مانند لحظه، پریود و عنصر زمانی را ارائه خواهیم کرد.در معرفی مدل شی‌گرا نیز ویژگی شی‌ها از دیدگاه استاندارد معرفی‌شده توسط گروه ODMG را بررسی می‌نماییم.

در فصل سوم دو مدل مشابه با مدل پیشنهادی ارائه کرده‌ایم که اولین آن مدل داده شی‌گرای زمانی نام دارد. در این مدل به بررسی چگونگی ادغام بعد زمان در عناصرداده‌ای خواهیم پرداخت. عمل درج و حذف یک شی با توجه به اصول زمان در مدل را بررسی می‌کنیم. مدل دوم، مدل داده ای یکپارچه زمانی است که یک مدل زمانی بر مبنای اصول رابطه‌ای می‌باشد. این مدل بر مبنای دو بعد از زمان- زمان اعتبار و زمان تراکنش- می‌باشد. در ادامه چند شمای رابطه‌ای در این زمینه معرفی می‌کنیم.

در فصل چهارم مدل پیشنهادی خود را که یک مدل شی‌گرای دوزمانی می‌باشد ارائه می‌کنیم. هشت روتین معرفی شده در این مدل قابلیت‌هایی را فراهم می‌آورد تا بواسطه آن از توانمندی‌های شی‌گرایی در استفاده از زمان به نحو مطلوبی عمل کند. لحاظ کردن زمان در شی و چگونگی کاربرد آن جهت پیاده‌سازی صحیح اصول مبتنی بر زمان بودن از مواردی می‌باشند که در ارائه این روتین بیش از هر مورد دیگری حائز اهمیت بوده است. در پایان این فصل نیز ارزیابی بین مدل پیشنهادی و مدلهای مطرح‌شده در فصل سوم و برخی دیگر از مدل‌ها ارائه خواهیم کرد.

در فصل پنجم نیز مروری بر تحقیقات صورت گرفته در زمینه مدل پیشنهادی و این پایان نامه خواهیم داشت و همچنین پیشنهاداتی مبتنی بر توسعه و تکامل مدل پیشنهادی و استفاده از این مدل در مفاهیم داده کاوی ارائه شده است.

فصل دوم
ادبیات تحقیق

۲-۱- مقدمه

در فصل پیشین به بررسی موارد مطالعه شده در این پایان نامه و چرایی ورود بدین مقوله پرداختیم. از اینرو در این فصل به بیان مفاهیم بنیادی که اهداف این پایان نامه بر آن بنا گردیده است و بدان‌ها نیازمند خواهیم بود، می-پردازیم. از آنجاییکه غایت نهایی، ارائه یک مدل داده‌ای مبتنی بر شی می‌باشد که بیشتر جنبه‌های زمان را نیز در خود می‌گنجاند، پس لازم می‌باشد تا مفاهیم کلی پایگاه داده شی‌گرا و پایگاه داده زمانی، نقاط قوت و ضعف هر یک را مطرح کنیم.

بنابراین در ابتدای این فصل تعریفی جامع و کامل از داده و اطلاع و موجودیت را ارائه خواهیم کرد. در ادامه به بررسی مدل‌های مختلف پایگاه دادگان و سیستم های مدیریت پایگاه دادگان می‌پردازیم و در نهایت هم پایگاه و مدل داده ای زمانی و شی‌گرا را بطور کامل معرفی خواهیم نمود.

۲-۲- داده، اطلاع و موجودیت

سه عنصر اصلی هر منبع داده‌ای عبارتند از داده، اطلاع و موجودیت. جهت درک بهتر مفاهیم در مدلی که در فصول بعدی ارائه می کنیم باید این سه عنصر کلیدی را از یکدیگر تمیز داد[17].

۲-۲-۱- داده

داده[1] دارای تعاریف مختلفی است، به طور کلی داده‌ها کلمات و ارزش‌های واقعی هستند که از طریق مشاهده و تحقیق بدست می آیند، به عبارت دیگر داده نمودی از وقایع، معلومات، رخدادها، پدیده‌ها و مفاهیم می‌باشد. در محاسبات، داده به اطلاعی گفته می شود که به شکلی مناسب برای انتقال و پردازش ترجمه شود. در کامپیوتر و رسانه‌های ارتباطاتی امروزی داده به شکل باینری تبدیل می شود. بنابراین داده یک نمایش باینری از یک موجودیت منطقی ذخیره شده در حافظه کامپیوتر است. از نظر ساختاری داده به مقادیر صفت خاصه انواع موجودیت‌ها اطلاق می‌شود. ریشه کلمه داده از عبارت لاتین Datum گرفته شده که به معنی اطلاع و Data فرم جمع آن است. اما Datum بندرت استفاده می‌شود و اکثرا ترجیح می‌دهند Data را به صورت یک کلمه مفرد استفاده کنند.

۲-۲-۲- اطلاع

اطلاع[2] مفهومی است که برای گیرنده آن قابل درک بوده و با دانستن آن می تواند برای حل یک مسئله تصمیم-گیری یا ارزیابی کند. وقتی اطلاع وارد کامپیوتر شده، ذخیره می گردد و به داده تبدیل می شود. بعد از پردازش، داده خروجی مجددا به عنوان اطلاع دریافت می‌شود. وقتی اسم صفت خاصه و مقدار منسوب به آن در دست باشند اطلاعی در مورد موجودیت حاصل می شود. هنگامیکه اطلاعات در یک بسته خاص قرار می‌گیرند و برای درک موضوعی یا انجام کاری استفاده می‌شود به دانش[3] تبدیل می‌شود .

۲-۲-۳- موجودیت

موجودیت[4] مفهوم کلی پدیده، شیء یا فردی است که در مورد آن می خواهیم اطلاع داشته باشیم و در کامپیوتر ذخیره کنیم. هر نوع موجودیت دارای مجموعه‌ای از صفات خاصه[5] است که ویژگی جداکننده یک نوع موجودیت از نوع دیگر محسوب می شود. به دو مثال زیر توجه فرمایید:

- اگر در نظر داریم یک سیستم پایگاه داده برای یک دبیرستان پیاده‌سازی کنیم مواردی چون دانش‌آموزان، دبیران، دروس، کلاس‌ها و غیره جزء موجودیت‌های سیستم به شمار می‌روند.
- موجودیت دانشجو در سیستم دانشگاه می‌تواند دارای صفات خاصه: نام، نام خانوادگی، سن، سال تولد، رشته تحصیلی، سال ورود و ... باشد و یا موجودیت درس صفات خاصه: کد درس، نام درس و تعداد واحد.

---

[1] Data
[2] Information
[3] Knowledge
[4] Entity
[5] Attribute

۲-۳- انواع دادگان ها

دادگان از نظر ساختار مفهومی و شیوه‌های رفتار با داده‌ها انواع مختلفی دارند که در ادامه هر یک را به صورت تفضیلی ارائه می نماییم.

۲-۳-۱- مدل های پایگاه داده

شگردهای مختلفی برای مدل‌های داده‌ای وجود دارد. بیشتر سیستم‌های پایگاه داده‌ای هر چند که بطور معمول بیشتر ازیک مدل را مورد حمایت قرار می‌دهند، حول یک مدل مشخص ایجاد شده‌اند. برای هر یک ازالگوهای منطقی[۱] اجراهای فیزیکی مختلفی قابل پیاده شدن است وسطوح کنترل مختلفی در انطباق فیزیکی برای کاربران مهیا می کند. یک انتخاب مناسب تاثیر موثری بر اجرا دارد.

یک مدل داده‌ای تنها شیوه ساختمان بندی داده‌ها نیست بلکه معمولا به صورت مجموعه‌ای از عملیات که می تواند روی داده‌ها اجرا شود تعریف می‌شوند. برای مثال در مدل رابطه‌ای، عملیاتی همچون گزینش[۲]، طرح‌ریزی[۳] و اتصال[۴] تعریف می گردد. در ادامه لیستی از این مدل‌ها را معرفی می‌نماییم[3].

- **مدل تخت**

مدل تخت یا جدولی[۵] تشکیل شده است از یک آرایه دو بعدی با عناصر داده‌ای که همه اجزای یک ستون به صورت داده‌های مشابه فرض می‌شودو همه عناصر یک سطر با هم در ارتباط هستند. برای نمونه در ستون‌هایی که برای نام کاربری و رمز عبور در جزئی از سیستم‌های پایگاه داده‌ای امنیتی مورد استفاده قرارمی گیرد هر سطر شامل رمز عبوری است که مخصوص یک کاربر خاص است. ستون های جدول که با آن در ارتباط هستند به صورت داده کاراکتری، اطلاعات زمانی، عدد صحیح یا اعداددممیز شناور تعریف می‌شوند. این مدل پایه برنامه‌های محاسباتی[۶] است. پایگاه داده ها با فایل‌های تخت به سادگی توسط فایل های متنی تعریف می‌شوند. هر رکورد یک خط است و فیلدها به کمک جدا کننده‌هایی از هم مجزا می‌شوند. داده‌های هر ستون مشابه هم است و ما به آن ستون‌ها فیلد ها[۷] می گوییم و هر خط را غیر از خط اول یک رکورد[۸] می‌نامیم. خط اول را که برخی پایگاه‌های داده‌ای آنرا ندارند رکورد برچسب[۹] گوییم. برای مثال به جدول ۲-۱ توجه فرمایید.

---

[1] logical model
[2] Selection
[3] Projection
[4] Join
[5] Flat(Table) model
[6] Spreadsheet
[7] Fields
[8] Record
[9] Field lables

جدول ۲-۱ نمونه اطلاع ها در مدل تخت

team	name	Id
Blues	Amy	۱
Reds	Bob	۲
Blues	Chuck	۳
Blues	Dick	٤
Reds	Ethel	٥
Blues	Fred	٦
Blues	Gilly	۷
Reds	Hank	۸

هر مقدار داده‌ای اندازه خاص خود را دارد که اگر به آن اندازه نرسد می‌توان ازکاراکتر فاصله برای این منظور استفاده کرد اما این مسئله مخصوصا زمانی که بخواهیم اطلاعات را بر روی کارت‌های منگنه قرار دهیم مشکل ساز خواهد شد. امروزه معمولا ازنویسه TAB برای جداسازی فیلد ها و کاراکتر خط بعد برای رکورد بعدی استفاده می‌کنیم. البته شیوه‌های دیگری هم وجود دارد مثلا به مثال زیر دقت کنید :

```
"1","Amy","Blues"
"2","Bob","Reds"
"3","Chuck","Blues"
"4","Dick","Blues"
"5","Ethel","Reds"
"6","Fred","Blues"
"7","Gilly","Blues"
"8","Hank","Reds"
```

این مثال از جدا کننده کاما استفاده می‌کند.در این نوع مدل تنها قابلیت حذف، اضافه، دیدن و ویرایش وجود دارد که ممکن است کافی نباشد. البته Microsoft Excel این مدل را پیاده سازی می‌کند.

- **مدل شبکه ای**[۱]

در سال ۱۹۶۹ و در کنفرانس زبان‌های سیستم‌های داده‌ای[۲] توسط Charles Bachman ارائه شد. در سال ۱۹۷۱ مجددا مطرح شد و اساس کار پایگاه داده‌ای قرار گرفت و در اوایل دهه ۸۰ با ثبت آن درسازمان بین‌المللی استانداردهای جهانی یا ISO به اوج رسید.

[1] Network
[2] CODASYL

مدل شبکه‌ای برپایه دو سازه مهم یعنی مجموعه‌ها و رکوردها ساخته می‌شود و برخلاف روش سلسله مراتبی که از درخت استفاده می‌کند، گراف را به کار می‌گیرد. مزیت این روش برسلسله مراتبی این است که مدل‌های طبیعی بیشتری را با قابلیت برقراری ارتباط بین موجودیت‌ها فراهم می‌کند. علیرغم این مزیت‌ها به دو دلیل اساسی این مدل با شکست مواجه شد: اول اینکه شرکت IBM با تولید محصولات IMS و DL/I که بر پایه مدل سلسله مراتبی است این مدل را نادیده گرفت. دوم اینکه سرانجام مدل رابطه‌ای جای آن راگرفت چون سطح بالاتر و واضح-تر بود. تا اوایل دهه ۸۰ به علت کارایی رابط‌های سطح پایین مدل سلسله مراتبی و شبکه‌ای پیشنهاد می‌شد که بسیاری از نیاز های آن زمان را برطرف می‌کرد. اما با سریعتر شدن سخت افزار به علت قابلیت انعطاف و سودمندی بیشترسیستم های رابطه‌ای به پیروزی رسیدند.

رکورد ها در این مدل شامل فیلدهایی است (ممکن است همچون زبان کوبول[1] به صورت سلسله مراتب اولویتی باشد). مجموعه‌ها با ارتباط یک به چند بین رکورد ها تعریف می‌شود: یک مالک و چند عضو. عملیات مدل شبکه‌ای از نوع هدایت‌کننده است: یک برنامه درموقعیت جاری خود باقی می‌ماند واز یک رکورد به رکورد دیگر می‌رود هرگاه که ارتباطی بین آنها وجود داشته باشد. معمولا از اشاره‌گرها[2] برای آدرس‌دهی مستقیم به یک رکورد در دیسک استفاده می‌شود. با این تکنیک کارایی بازیابی اضافه می‌شود هر چند در نمایش ظاهری این مدل ضروری نیست.

- **مدل رابطه ای**

مدل رابطه‌ای در یک مقاله تحصیلی توسط E. F. Codd در سال ۱۹۷۰ ارائه گشت. این مدل یک مدل ریاضیاتی است که با مفاهیمی چون مستندات منطقی[3] و تئوری مجموعه‌ها[4] در ارتباط است. محصولاتی همچون اینگرس[5]، اراکل[6]،DB2,SQL Server براین پایه ایجاد شده است. ساختار داده ها در این محصولات به صورت جدول است با این تفاوت که می‌تواند چند سطر داشته باشد. به عبارت دیگر دارای جداول چند گانه است که به طور صریح ارتباطات بین آنها بیان نمی‌شود و در عوض کلیدهایی به منظور تطبیق سطرها در جداول مختلف استفاده

---

[1] COBOL
[2] Pointers
[3] Predicate logic
[4] Set theory
[5] Ingress
[6] Oracle

می‌شود. به عنوان مثال جدول کارمندان ممکن است ستونی به نام "موقعیت" داشته باشد که کلید جدول موقعیت را با هم تطبیق می‌دهد.

- **پایگاه داده‌های چند بعدی**

پایگاه داده‌های رابطه‌ای توانست به سرعت بازار راتسخیر کند، هرچند کارهایی نیز وجود داشت که این پایگاه داده‌ها نمی‌توانست به خوبی انجام دهد. به ویژه بکارگیری کلیدها در چند رکورد مرتبط به هم و در چند پایگاه داده مشترک، کندی سیستم را موجب می‌شد. برای نمونه برای یافتن نشانی کاربری با نام دیوید، سیستم رابطه‌ای باید نام وی را در جدول کاربر جستجو کند و کلید اصلی[1] را بیابد و سپس در جدول نشانی‌ها، دنبال آن کلید بگردد. اگر چه این وضعیت از نظر کاربر، فقط یک عملیات محسوب، اما به جستجو درجداول نیازمند است که این-کار پیچیده و زمان‌بر خواهد بود. راهکار این مشکل این است که پایگاه داده‌ها اطلاعات صریح درباره ارتباط بین داده‌ها را ذخیره نماید. می‌توان به جای یافتن نشانی دیوید با جستجوی کلید در جدول نشانی، اشاره‌گر به داده‌ها را ذخیره نمود. درواقع، اگر رکورد اصلی، مالک داده باشد، در همان مکان فیزیکی ذخیره خواهد شد و ازسوی دیگر سرعت دسترسی افزایش خواهد یافت.

چنین سیستمی را پایگاه داده‌های چندبعدی می‌نامند. این سیستم در هنگامی که از مجموعه داده‌های بزرگ استفاده می‌شود، بسیار سودمند خواهد بود. از آنجایی که این سیستم برای مجموعه داده‌های بزرگ به کارمی‌رود، هیچگاه در بازار به طور مستقیم عمومیت نیافتند.

- **پایگاه داده‌های شیء**

اگر چه سیستم‌های چند بعدی نتوانستند بازار را تسخیر نمایند، اما به توسعه سیستم‌های شیء منجر شدند. این سیستم‌ها که مبتنی بر ساختار و مفاهیم سیستم‌های چند بعدی هستند، به کاربر امکان می‌دهند تا اشیاء را به طور مستقیم درپایگاه داده‌ها ذخیره نماید. بدین ترتیب ساختار برنامه نویسی شیءگرا[2] را می‌توان به طور مستقیم و بدون تبدیل نمودن به سایر فرمت‌ها، درپایگاه داده‌ها مورد استفاده قرار داد. این وضعیت به دلیل مفاهیم مالکیت[3] در سیستم چندبعدی، رخ می‌دهد. در برنامه شیء گرا[4]، یک شیء خاص" مالک " سایر اشیاء در حافظه است، مثلا دیوید مالک

[1] Primary key
[2] Object Oriented
[3] Ownership
[4] Object Oriented Programming(OOP)

نشانی خود می‌باشد. در صورتی که مفهوم مالکیت در پایگاه داده‌های رابطه‌ای وجود ندارد. در ادامه بطور کامل مفاهیم پایگاه داده‌ای شی‌گرا را بررسی خواهیم کرد.

۲-۳-۲- سیستم مدیریت پایگاه دادگان[1]

برنامه رایانه‌ای که برای مدیریت و پرسش و پاسخ بین پایگاه‌های داده‌ای استفاده می‌شود را مدیر سیستم پایگاه داده‌ای یا به اختصار DBMS می‌نامیم. خصوصیات و طراحی سیستم‌های پایگاه داده‌ای در علم اطلاعات مطالعه می‌شود.

مفهوم اصلی پایگاه داده این است که پایگاه داده مجموعه‌ای از رکوردها یا تکه‌هایی از یک شناخت است. نوعا در یک پایگاه داده توصیف ساخت یافته‌ای برای موجودیت‌های نگه‌داری شده در پایگاه داده وجود دارد: این توصیف با یک الگو یا مدل شناخته می‌شود. مدل توصیفی، اشیا پایگاه‌های داده و ارتباط بین آنها را نشان می‌دهد. روش‌های متفاوتی برای سازماندهی این مدل‌ها وجود دارد که به آنها مدل‌های پایگاه داده گوییم. پرکاربردترین مدلی که امروزه بسیار استفاده می‌شود، مدل رابطه‌ای است که به طور عام بدین صورت تعریف می‌شود: نمایش تمام اطلاعاتی که به فرم جداول مرتبط که هریک از سطرها و ستونها تشکیل شده است(تعریف حقیقی آن در علم ریاضیات بررسی می‌شود). در این مدل وابستگی‌ها به کمک مقادیر مشترک در بیش از یک جدول نشان داده می‌شود. مدل‌های دیگری مثل مدل سلسله مراتب و مدل شبکه‌ای به طور صریح تری ارتباط‌ها را نشان می‌دهند.

در مباحث تخصصی تر اصطلاح دادگان یا پایگاه داده به صورت مجموعه‌ای از رکوردهای مرتبط با هم تعریف می‌شود. بسیاری از افراد خبره مجموعه‌ای از داده‌هایی با خصوصیات یکسان به منظور ایجاد یک پایگاه داده‌ای یکتا استفاده می‌کنند. معمولاً DBMS ها بر اساس مدل‌هایی که استفاده می‌کنند تقسیم بندی می‌شوند: ارتباطی، شی‌گرا، شبکه‌ای و امثال آن. مدل‌های داده‌ای به تعیین زبان‌های دسترسی به پایگاه‌های داده علاقه‌مند هستند. بخش قابل توجهی از مهندسی DBMS مستقل از مدل‌ها می‌باشد و به فاکتورهایی همچون اجرا، هم‌زمانی، جامعیت و بازیافت از خطاهای سخت افزاری وابسته است. در این سطح تفاوت‌های بسیاری بین محصولات وجود دارد.

ویژگی‌های سیستم مدیریت پایگاه داده‌ها

اکنون پس از این بررسی به توصیف سیستم مدیریت پایگاه داده‌ها می‌پردازیم. سیستم مدیریت پایگاه داده‌ها، مجموعه‌ای پیچیده از برنامه‌های نرم‌افزاری است که ذخیره سازی و بازیابی داده‌های (فیلدها، رکوردها و فایل‌ها) سازمان را در پایگاه داده‌ها، کنترل می‌کند. این سیستم، کنترل امنیت و صحت پایگاه داده‌ها را نیز بر عهده دارد.

---

[1] DBMS