

UDC: 004.622

INFO M: str. 35-40

**IMPLIKACIJE VREMENA U PROJEKTOVANJU SKLADIŠTA  
PODATAKA KOJE PODRŽAVA CRM  
THE IMPLICATIONS OF TIME IN DESIGNING A DATA  
WAREHOUSE THAT SUPPORTS CRM**

**Dragana Čamilović**

**REZIME:** Bez prisustva CRM-a skladišta podataka ne bi bila toliko popularna. CRM je strategija za optimizaciju doživotne vrednosti kupca kroz njegovo bolje upoznavanje. Ovo je omogućeno zahvaljujući skladištu podataka, ali je vrlo važno da ono arhivira tačne podatke o kupcima i njihovom ponašanju. Stoga se skladišta podataka moraju projektovati tako da se uzme u obzir vreme.

**KLJUČNE REČI:** skladište podataka, CRM, uloga vremena

**ABSTRACT:** Without CRM data warehouses would not be so popular. CRM is a strategy for optimizing the lifetime value of customer by getting to know the customer better. A data warehouse makes this possible, but it is very important that it archives accurate data about customers and their behavior. Therefore data warehouses have to be designed to take account of time.

**KEY WORDS:** data warehouse, CRM, the role of time

## 1. UVOD

Interesovanje za projektovanje i implementaciju skladišta podataka bi se smanjilo da na scenu nije stupila relativno mlada naučna disciplina poznata pod nazivom upravljanje odnosima sa kupcima, odnosno korisnicima usluga (CRM). CRM je usmeren na stvaranje, razvijanje i jačanje dugoročnih veza sa kupcima/korisnicima, kako bi se optimizirala njihova doživotna vrednost. CRM je filozofija koja treba da omogućiti potpunu orijentaciju na kupca/korisnika i stoga pretpostavlja postojanje izvora podataka, kako o samim kupcima/korisnicima (demografski podaci), tako i o njihovom ponašanju (bihevioristički podaci). Stoga je sa stanovišta CRM-a jako važno da skladište poseduje tačne podatke o kupcima/korisnicima, te se mora pratiti na koji se način oni tokom vremena menjaju. Ovaj će se rad baviti upravo ovim problemom. Ali, najpre treba dati odgovor na pitanje: šta je to skladište podataka? O tome će biti reči u narednom odeljku rada.

## 2. ODREĐENJE POJMA SKLADIŠTA PODATAKA

Skladište podataka treba da čuva podatke koji potiču iz različitih operacionih, eksternih i ostalih baza podataka. Ono predstavlja centralni izvor podataka koji su prethodno pročišćeni, transformisani i katalogizirani, kako bi služili za potrebe podrške donošenju poslovnih odluka [1]. Skladište podataka se, dakle, implementira kako bi se najznačajnije poslovne informacije stavile u ruke donosiocima odluka [2].

Inmon smatra da skladište podataka predstavlja veliko, jedinstveno, integrisano, prilagodljivo, elastično i bezbedno spremište podataka koje obezbeđuje infrastrukturnu osnovu informacionih softverskih aplikacija u organizaciji [3]. On ističe da su njegove najznačajnije osobine [4, 5]:

- orijentisanost na predmet,
- integrisanost,
- vremenska dimenzija i
- nepromenljivost.

Dok su klasični transakcioni sistemi usmereni ka funkcionalnim aplikacijama, skladišta podataka su orijentisana na predmete. Pri tome svaka kompanija poseduje sopstveni skup predmeta, koji je pre svega definisan delatnošću kojom se ona bavi (npr. preduzeće koje se bavi maloprodajom posmatra jedan skup predmeta, osiguravajuća kompanija drugi itd.).

Već na samom početku ovog odeljka je pomenuto da skladište treba da integriše podatke iz različitih izvora. Pri tome jedan isti atribut može imati različite vrednosti u različitim aplikacijama. Dalje, mogu se upotrebljavati različite merne jedinice, ključevi mogu biti različitog tipa, a često ne postoji konzistentnost ni u nazivima [4]. Zato je vrlo važno integrisati podatke pre nego što se unesu u skladište.

Vremenska dimenzija podrazumeva da se podaci u skladištu obično vezuju za određeni vremenski period (npr. 5 do 10 godina) [6]. Svaki ključ poseduje, bilo eksplicitno ili implicitno, vremensku dimenziju [4, 6], tj. svaki podatak je validan u određenom vremenskom periodu. Zato Inmon ovu karakteristiku ističe kao jednu od najznačajnijih osobina skladišta podataka.

Za razliku od transakcionih sistema, gde su operacije ažuriranja česte, skladište podataka je nepromenljivo tj. ovde se ne vrši ažuriranje. Sa stanovišta skladišta su obično podržane dve operacije nad podacima: unos podataka i pristup podacima [6]. Podaci se pri tome ne unose slog po slog, već se periodično unosi veća količina podataka odjednom. Oni se nakon toga ne menjaju, već se njima samo po potrebi može pristupati radi vršenja različitih analiza.

Na osnovu svega izloženog je jasno da skladište podataka treba da donosiocima odluka stavi na raspolaganje velike količine istorijskih podataka prikupljenih, pre svega, iz operativnih sistema namenjenih transakcionoj obradi. Nad ovim podacima se mogu dalje vršiti najrazličitije analize, a sve u cilju unapređenja procesa odlučivanja na nivou čitave kompanije (između ostalog i unapređenja CRM-a). Dakle, skladište podataka predstavlja jedan savremeni sistem za podršku odlučivanju, zasnovan na podacima.

Jedan od pristupa projektovanju skladišta podataka se zasniva na tzv. dimenzionom modelu. Kada su počela da se grade skladišta uočeno je da korisnici (a to su pre svega donosioci odluka) žele da analiziraju podatke po većem broju dimenzija. Tako je Codd prvi primetio da višedimenzioni konceptijski pogled za većinu poslovnih ljudi predstavlja prirodan način za posmatranje njihovih preduzeća [7]. Dimenziona analiza je tako postala metod za definisanje skladišta podataka [7]. O ovome će više reči biti u narednom odeljku rada.

### 3. DIMENZIONA ANALIZA

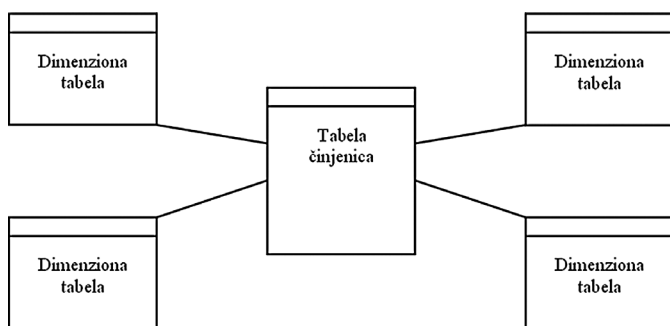
Dimenziona analiza treba da na osnovu razgovora sa donosiocima odluka utvrdi koji ih predmeti najviše zanimaju i koje su najvažnije dimenzije analize [7]. Na taj način oni definišu mere i dimenzije.

Dimenzioni model se lako prevodi u relacioni model, pa je tako moguće napraviti relacionu šemu i to zvezdastu (star schema), ili pahuljičastu (snowflake schema).

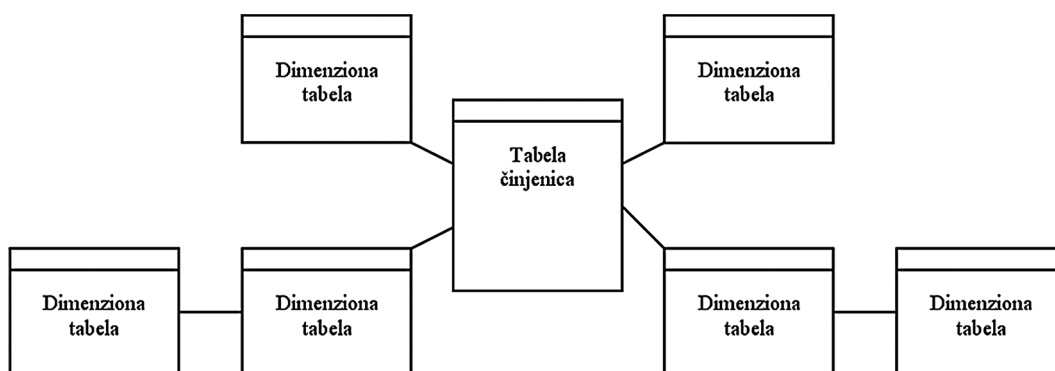
Model treba da ima u vidu potrebe korisnika i nivo detaljnosti podataka koji će biti neophodan za vršenje analiza. Njime se definiše i koje će se mere i dimenzije posmatrati. Mora se voditi računa da odabrano rešenje podržava lako održavanje i eventualna unapređenja u budućnosti, shodno novim zahtevima korisnika.

Svaki model ima u svom središtu tabelu činjenica, koja omogućuje definisanje mera kocke podataka. Sa njom su povezane dimenzione tabele, a način na koji je to učinjeno zavisi od odabrane šeme.

Najjednostavniji model povezivanja je u okviru zvezdaste šeme. Centralno mesto u ovoj šemi ima tabela činjenica



Slika 1. – Zvezdasta šema



Slika 2. – Pahuljičasta šema

koja poseduje složen ključ, koji će omogućiti povezivanje sa dimenzionim tabelama, ali poseduje i po jedno numeričko polje za svaku meru koja će biti uključena u analizu. Svako dimenziji odgovara po jedna dimenziona tabela, koja je direktno povezana sa već pomenutom tabelom činjenica. Ovo je prikazano na Slici 1.

Ukoliko postoji dekompozicija jedne ili više dimenzija onda je reč o pahuljičastoj šemi. Pahuljičasta šema, zapravo, predstavlja oblik zvezdaste šeme kod koje je izvršena normalizacija pojedinih dimenzionih tabela, pa su podaci o nekoj dimenziji smešteni unutar većeg broja tabela. Ovaj je model razumljiv i prihvatljiv projektantima baza podataka, ali je zato poslovnim korisnicima teže da rade sa pahuljičastom, nego sa zvezdastom šemom (jer zvezdasta šema poseduje jednostavniju strukturu). Pahuljičasta šema može obezbediti uštedu memorijskog prostora [8], a zbog smanjenja redudanse podataka postiže se i lakše održavanje. Ipak treba imati u vidu da tabela činjenica po pravilu zauzima mnogo više memorije od dimenzionih tabela, pa ove uštede često nisu značajne.

Još jedan razlog zašto je pahuljičasta šema manje “popularna” je taj što izvršavanje upita zahteva veći broj spajanja (join), čime se postižu lošije performanse u odnosu na zvezdastu šemu [6]. Primer pahuljičaste šeme je prikazan na Slici 2.

### 4. VREMENSKA DIMENZIJA PRI PROJEKTOVANJU SKLADIŠTA PODATAKA KOJE PODRŽAVA CRM

Iako koncept izgradnje skladišta podataka ne predstavlja novinu, jer se o ovoj temi piše i govori poslednjih desetak godina, treba imati u vidu da unapređenje načina poslovanja kompanija zahteva i promene u načinu razmišljanja o projektovanju i implementaciji skladišta podataka koje treba da podrži donošenje poslovnih odluka. CRM pogotovo zahteva raspoloživost tačnih i kvalitetnih informacija, jer je nemoguće baviti se analizom kupaca/korisnika usluga, bez posedovanja informacija o njima samima i njihovom ponašanju. I upravo zato se implementacija skladišta podataka nametnula kao potreba savremenih preduzeća koja žele da budu usmerena na kupca/korisnika usluga.

Skladište podataka, kao što je ranije pomenuto, treba da, između ostalih, sadrži i demografske podatke o kupcima/korisnicima, ali i podatke o njihovom ponašanju, pri čemu se pod

ovim pojmom najčešće podrazumevaju podaci o načinjenim transakcijama (kupovinama ili načinu korišćenja usluga). Sa stanovišta CRM-a je vrlo važno napraviti razliku između ova dva tipa podataka.

Klasična skladišta podataka, ili, kako ih Chris Todman naziva, skladišta podataka prve generacije [7], uglavnom se fokusiraju na praćenje ponašanja kupaca/korisnika. To znači da je tabela činjenica, bilo da se radi o zvezdastoj ili pahuljičastoj šemi, po pravilu tabela koja sadrži podatke o ponašanju kupaca/korisnika usluga. To može biti tabela o kupovinama koje je kupac načinio, u slučaju telefonskih kompanija to je tabela koja sadrži podatke o načinjenim pozivima, dok hoteli mogu pratiti podatke o boravcima svojih klijenata, banke podatke o novčanim transakcijama, avio-kompanije se mogu fokusirati na podatke o rezervacijama itd. Tabela sa podacima o kupcima/korisnicima usluga obično se javlja kao dimenziona tabela u okviru zvezdaste ili pahuljičaste šeme. Ovakav pristup je pogodan za vršenje pojedinih analiza, ali on ne zadovoljava u potpunosti potrebe CRM-a. Ono što je sa aspekta CRM-a značajno jeste posmatranje promena podataka o kupcima/korisnicima. Ovo je pogotovo bitno za procenu stope gubitka klijenata, a njihov odliv je veliki problem za mobilne operatore, Internet provajdere, maloprodajna preduzeća, pa čak i problem banaka. Podaci o ponašanju kupaca/korisnika usluga (tzv. bihevioristički podaci) neće ukazati na razloge gubitka klijenata. Recimo, korisnik usluga mobilnog operatora može odlučiti da ode kod drugog provajdera jednostavno iz razloga promene mesta stanovanja. Možda u tom drugom mestu ne postoji dobra pokrivenost postojećeg operatora, ili se korisnik preselio u drugu državu pa ne želi da plaća roming. Dakle, sa stanovišta procene odliva korisnika za telekomunikacionu kompaniju je bitno da posmatra promenu adrese stanovanja. Zato Todman predlaže da se podaci o kupcima podele u dve kategorije [7]:

- Podaci koji se ne menjaju, ili njihova promena nije za nas interesantna.
- Podaci čiju promenu treba evidentirati u okviru skladišta podataka. Oni će, za potrebe izrade konceptualnog modela, biti nazvani promenljivim atributima.

Todman u konceptualni model skladišta podataka koje je orijentisano na kupca/korisnika usluga uključuje i tzv. izvedeni segment kupca/korisnika usluga, kao što je npr. procenjena ukupna vrednost kupca/korisnika tokom njegovog životnog veka ili sklonost ka odlivu. Naravno, nikako ne treba isključiti iz modela podatke o ponašanju kupaca/korisnika usluga.

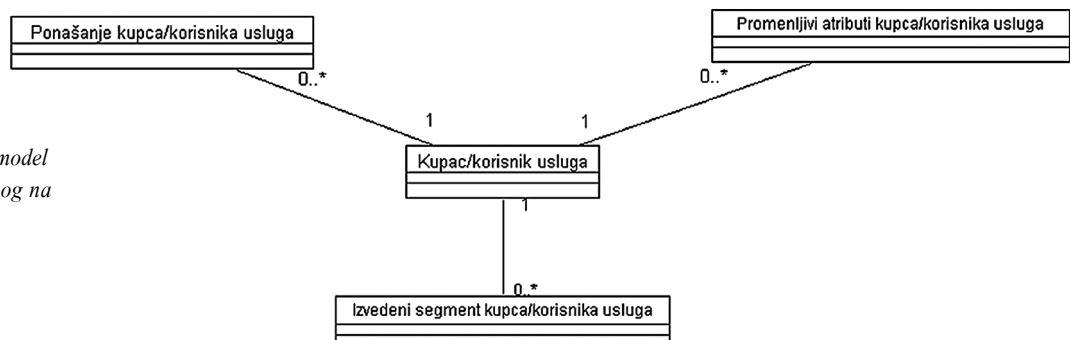
Koristeći UML notaciju, ovaj opšti konceptualni model, predložen u [7], može se predstaviti dijagramom klasa koji, izrađen u Rational Rose-u, izgleda kao na Slici 3.

Primetimo da je kardinalnost veza između tabele sa podacima o kupcima/korisnicima usluga i pomenutih tabela (sa podacima o promenljivim atributima i izvedenim segmentima kupca/korisnika) tipa nula ili više, što znači da možemo posmatrati i više promenljivih atributa, kao i više izvedenih segmenta kupca/korisnika usluga. Naravno, za svakog kupca/korisnika je karakteristično da se njegovo ponašanje može pratiti kroz čitav niz slogova, pa nije potrebno davati neko posebno objašnjenje ove kardinalnosti.

Tabela činjenica, koja sadrži podatke o ponašanju kupaca/korisnika usluga, obično poseduje i polje koje sadrži podatak o vremenu (s obzirom da se transakcije uvek prate u vremenu, a ova tabela najčešće sadrži podatke o transakcijama). Međutim, u OLTP sistemima se najčešće čuvaju samo podaci o ponašanju kupaca/korisnika generisani u poslednjih nekoliko dana ili meseci, jer je količina ovih podataka ogromna. Jedna od osnovnih razlika između OLTP sistema i skladišta podataka i leži u činjenici da skladišta treba da prikazuju informacije o tzv. istorijskim podacima, pa je samim tim vremenska dimenzija i te kako značajna za posmatranje. Nemogućnost OLTP sistema da posmatraju istorijske podatke i jeste dovela do potrebe za razvojem skladišta podataka. Stoga se u OLTP sistemima ne čuvaju svi podaci koji se čuvaju u skladištima. I pre nego što oni izbrišu iz transakcionih sistema, potrebno je, naravno, da već budu smešteni u skladište podataka (što i nije problem, jer se to može postići "bulk" obradom). Pored posmatranja vremenske dimenzije generisanja transakcionih podataka, potrebno je pratiti i kako se promenljivi atributi menjaju tokom vremena.

Jensen i dr. u [9] prave razliku između stvarnog vremena (valid time) i transakcionog vremena (transaction time) promene atributa. Stvarno vreme je onaj vremenski trenutak u kome zaista, u realnosti, promena nastaje. Transakciono vreme je, sa druge strane, onaj vremenski trenutak kada se ova promena registruje u bazi. Da li postoji razlika između ova dva vremena? O tome će biti reči u nastavku odeljka.

Polje u okviru tabele činjenica koje čuva podatak o vremenu (npr. vreme kupovine, porudžbine, u slučaju telekomunikacionih kompanija od značaja je posmatranje vremena kada je načinjen telefonski poziv, i slično) u velikom broju slučajeva predstavlja stvarno vreme (npr. pri evidentiranju telefonskih poziva), ali nekada je to ipak transakciono vreme tj. vreme promene atributa i trenutak kada je ta promena evidentirana unutar baze se razlikuju. S obzirom da, kako je zaključeno, treba pratiti i promenu pojedinih podataka o



Slika 3. – Opšti konceptualni model skladišta podataka usredsređenog na kupca/korisnika usluga

kupcima/korisnicima usluga, postavlja se pitanje kako ovde tretirati problem vremena? Kada dođe do promene atributa dimenzionih tabela vreme promene se ne čuva u polju unutar tabele. Razlog ovome je činjenica da sa stanovišta OLTP sistema vreme nastanka promene vrednosti atributa i nije značajno za posmatranje. Ako se npr. kupac/korisnik preseli, njegova će adresa biti zamenjena u bazi, pri čemu se stara po pravilu neće ni čuvati (tj. biće izvršena jednostavna update naredba). Čak je i pitanje kada će se ova promena u bazi izvršiti, i sigurno će postojati razlika između stvarnog i transakcionog vremena. Sa druge strane, za skladište podataka promena adrese jeste od značaja. Ako se npr. prate podaci o prodaji po pojedinim gradovima, onda je činjenica da se kupac preselio iz jednog grada u drugi i te kako važna pri analizi podataka. I, naravno, važno je znati i kada se to desilo tj. kada je kupac promenio adresu stanovanja.

Zbog velikog značaja posmatranja vremena sa stanovišta skladišta podataka Kimball predlaže da se ono uvede kao nova dimenziona tabela [10]. Može se postaviti pitanje zašto uvesti novu tabelu, ako SQL podržava brojne funkcije za rad sa vremenom, a već postoji odgovarajuće polje (tipa datetime ili time) unutar tabele činjenica? Kimball ističe da je svaka ispostava podataka (data mart), a koja predstavlja deo skladišta podataka namenjen da se njime služi određena organizaciona jedinica, kvalitetna koliko i njegove dimenzione tabele. Ukoliko se ne uvedu odgovarajuće dimenzije, nije moguće generisati sve potrebne izveštaje [10]. Tako novouvedena dimenziona tabela može sadržati i podatke o fiskalnim periodima i sl. [10], tj. u ovu tabelu se po potrebi mogu dodavati novi atributi, a svi ovi podaci ne bi bili dostupni ukoliko bi se umesto posebne tabele za praćenje vremena zadovoljili postojanjem polja u okviru tabele činjenica.

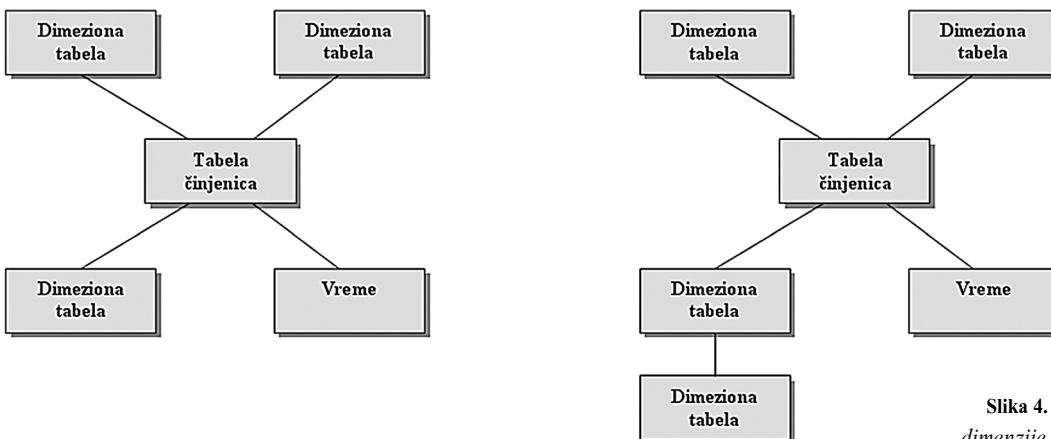
Kako i promene vrednosti pojedinih atributa dimenzionih tabela (a pre svega tabele o kupcima/korisnicima usluga) treba pratiti u vremenu, postavlja se pitanje da li to znači da treba dozvoliti povezivanje tabele sa podacima o vremenu sa dimenzionim tabelama? Kimball daje odričan odgovor na ovo pitanje, jer smatra da se semantika vremenske dimenzije kod dimenzionih tabela razlikuje u odnosu na tabelu činjenica [7]. Zvezdasta i pahuljičasta šema stoga dozvoljavaju samo povezivanje novouvedene dimenzije sa tabelom činjenica, kako je to prikazano na Slici 4.

Tabela Vreme mora da sadrži sve vremenske trenutke koji se pojavljuju kao vrednost odgovarajućeg polja u tabeli činjenica.

Dakle, svaki slog iz tabele činjenica je povezan sa slogom u tabeli Vreme i u najvećem broju slučajeva, kao što je više puta istaknuto, radi se o stvarnom vremenu. U slučaju promene vrednosti atributa unutar dimenzionih tabela obično je nemoguće znati stvarno vreme njihovog nastanka, ali se zato može zabeležiti njihovo transakciono vreme. Na koji način se ono može odrediti? Mogu se uvesti određeni trigeri koji će biti pokrenuti pri promeni vrednosti atributa u bazi. Drugi način, koji je najrasprostranjeniji, zasniva se na poređenju file-ova [7]. Prilikom dodavanja novih podataka u skladište vrši se poređenje sa podacima koji se već u njemu nalaze i svaka promena se evidentira. Ovaj pristup je naročito popularan zbog toga što ne zahteva bilo kakve promene unutar OLTP sistema, a pored toga značajna je i činjenica da je njegova realizacija prilično jeftina [7]. Sa druge strane, ovaj pristup ima i određene nedostatke. Neke kompanije mogu imati po nekoliko miliona kupaca/korisnika, pa je poređenje file-ova jako vremenski zahtevno. U tom slučaju uputno je vršiti poređenje file-ova vikendima, i ne toliko često (npr. jednom mesečno) [7]. Naravno, što se rede operacija poređenja vrši, to je diskutabilnija tačnost podataka sačuvanih u skladištu.

Na koji način pratiti promene atributa unutar dimenzionih tabela? Odgovor na ovo pitanje zavisi od toga da li su promene pojedinih atributa u tim tabelama učestale ili ne. Tako je Kimball još 1994. godine uveo pojam dimenzija koje se sporo menjaju [11, 12, 13]. Ovo su dimenzije koje su gotovo konstantne u vremenu [13] i često se u literaturi pominju kao SCDs (Slowly Changing Dimensions). Za svaki atribut u okviru dimenzione tabele treba definisati strategiju praćenja njegove promene tokom vremena. Kimball predlaže tri načina praćenja promene vrednosti atributa, tj. sledeće tri tehnike: upisivanje nove vrednosti atributa umesto stare, kreiranje novog sloga u okviru dimenzione tabele i kreiranje polja koje će sadržati novu vrednost atributa [12, 13]. Prva dva koncepta su interesantna sa aspekta projektovanja skladišta podataka za potrebe CRM-a, ali ovde će biti dat kratak prikaz sve tri strategije.

Već je pri razmatranju predloženog konceptualnog modela napravljena razlika između tzv. promenljivih i nepromenljivih atributa. Nepromenljivi atributi, zapravo, mogu menjati vrednosti tokom vremena, ali to nije interesantno sa aspekta projektovanja skladišta podataka. Sa druge strane možemo evidentirati novu vrednost atributa, ali ne moramo nigde čuvati staru (npr. ukoliko su pogrešno uneseni nazivi gradova



Slika 4. – Uvođenje vremena kao nove dimenzije i povezivanje sa tabelom činjenica

stanovanja kupaca/korisnika usluga, uglavnom je dovoljno da stara vrednost bude zamenjena novom [12]). U ovom slučaju možemo primeniti prvi pristup koji Kimball predlaže, koji podrazumeva da se stara vrednost atributa jednostavno zameni novom. Ovo je daleko najjednostavniji način rešavanja problema promene vrednosti atributa. Osim toga je i brz [13]. Međutim, osnovni nedostatak je u tome što se gubi mogućnost praćenja promene atributa tokom vremena, pa ga je uglavnom preporučljivo koristiti u slučajevima kada su potrebne ispravke pogrešno unesenih podataka. Na žalost, mnogi projektanti skladišta podataka često koriste ovaj pristup i kada je njegova upotreba potpuno nepodesna.

Nekada, međutim, nije podesno izvršiti samo zamenu stare vrednosti atributa novom. Već je razmatrana problematika evidentiranja promene adrese kupca/korisnika usluge. Ovde treba primeniti drugopomenutu strategiju. Ona predlaže da se svaki put kada dođe do promene vrednosti atributa kupca/korisnika usluga u skladište unosi nov slog u tabelu o kupcima/korisnicima, koji će naravno sadržati tu novu vrednost atributa, dok će vrednosti svih ostalih atributa ostati nepromenjene. Novi slog ne sme imati istu vrednost ključa (jer u tom slučaju dodavanje ovog sloga ne bi ni bilo moguće) i zato se koriste tzv. ključevi surogati. Ključ surogat obično se kreira tako što se na postojeću vrednost ključa posmatranog kupca/korisnika usluga dodaje broj koji ukazuje na novu verziju sloga tog istog kupca/korisnika [12]. Kupac čija je vrednost ključa 1234 bi tako dobio vrednost ključa 1234001. Nakon prve promene vrednosti atributa ubacio bi se nov slog sa vrednošću ključa 1234002 itd. Obično se radi o celobrojnim vrednostima, tj. podacima tipa integer koji zauzimaju svega četiri bajta [14]. Druga opcija jeste korišćenje složenog ključa, kao što je prikazano u Tabeli 1. [7].

ID kupca/korisnika usluge	Broj "verzije" sloga
1234	001
1234	002
1234	003

Tabela 1. – Složeni ključ surogat

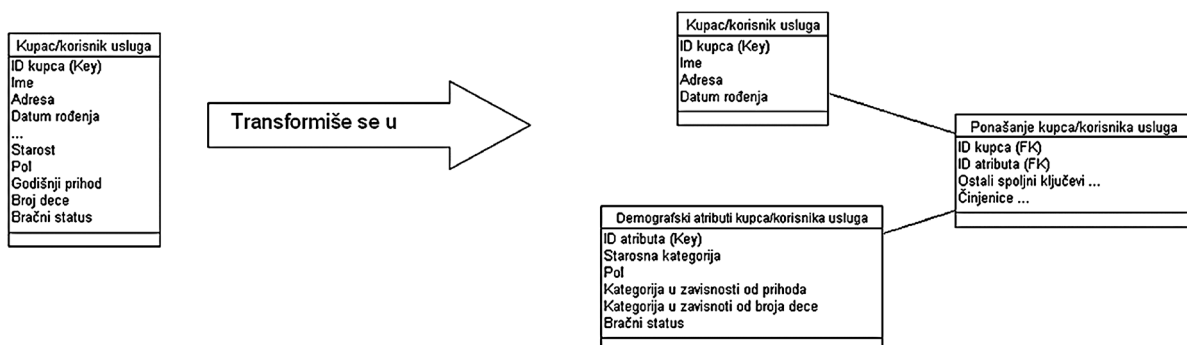
Administrator skladišta podataka snosi odgovornost za kontrolu dodeljivanja vrednosti ključeva, kao i održavanje broja verzije sloga [7].

Treći način rešavanja promene vrednosti atributa je pogodan samo u slučajevima kada je potrebno posmatrati tzv. "alternativnu realnost" (neki autori koriste ovaj termin da bi

ukazali na činjenicu da je potrebno posmatrati i staru i novu vrednost atributa), što je uglavnom slučaj sa onim atributima čija je promena vrednosti rezultat ljudske odluke [13]. Ovaj metod podrazumeva dodavanje novog polja u okviru dimenzione tabele, pa će nam se na raspolaganju, očigledno, nalaziti i nova i stara vrednost atributa. Time je omogućeno da se podaci iz tabele činjenica razmatraju po obe vrednosti atributa, pa ovaj pristup isključivo treba koristiti kada postoji izričit zahtev za ovakvom vrstom uporednih analiza. Zato se on u slučaju projektovanja skladišta podataka koje treba da podrži CRM jako retko koristi.

Sve pomenute strategije su se odnosile na dimenzije koje se sporo menjaju tj. SCDs. Iako je drugopomenut pristup najrasprostranjeniji, on daje dobre rezultate samo ukoliko dimenziona tabela nije suviše velika i ukoliko promene vrednosti atributa u njoj nisu previše česte. Firme koje imaju manje od 100000 kupaca/korisnika u okviru dimenzione tabele u kojima se čuvaju podaci o njima bez problema mogu ubacivati nove slogove svaki put kada dođe do promene vrednosti nekog od pomenutih promenljivih atributa. Ali šta da rade oni koji imaju na milione kupaca/korisnika? Kimball predlaže korišćenje tzv. minidimenzija. Tako se npr. može formirati minidimenzija u kojoj će se čuvati demografski atributi kao što su starost, pol, broj dece (ili članova porodice), kategorija u zavisnosti od godišnjeg prihoda i slično [13]. U minidimenziji će se za svaku kombinaciju vrednosti ovih atributa uneti po jedan slog, što znači da se ne unosi slog za svakog kupca/korisnika usluga [13]. Kompromis koji je pri ovom pristupu nužno napraviti je sledeći: umesto posmatranja svih mogućih vrednosti pojedinih atributa treba formirati određene kategorije u zavisnosti od opsega vrednosti atributa. Npr. neće biti posmatrane sve moguće vrednosti atributa (polja) Starost, već će se posmatrati starosne kategorije. Sličan princip se može primeniti i na posmatranje kategorija na osnovu godišnjih prihoda kupca/korisnika usluga. Uvođenje kategorija znatno smanjuje broj kombinacija koje se čuvaju u okviru minidimenzije. Međutim, vrlo je važno dobro razmisliti na koji način će se one formirati, jer se kasnije ne praktikuje njihova izmena. Tako se, dakle, originalna dimenziona tabela transformiše u dve tabele na način kako je to prikazano na Slici 5.

Po potrebi se može uvesti i više od jedne minidimenzije. Korišćenje minidimenzija je, svakako, rešenje za kompanije koje imaju jako veliki broj kupaca/korisnika usluga, tj. u slučaju postojanja velikih dimenzija koje se brzo menjaju.



Slika 1. – Uvođenje minidimenzije [13]

## 5. ZAKLJUČAK

Vreme ima veliki uticaj na skladišta podataka, jer su ona u krajnjoj liniji vremenske aplikacije. Međutim, tek je pojava CRM-a istakla problem predstavljanja vremena, jer odsustvo posmatranja implikacije vremena bi skladišta učinila neupotrebljivim sa stanovišta CRM-a.

U ovom radu su posmatrani samo neki aspekti implikacije vremena na projektovanje i implementaciju skladišta podataka, koje je autor smatrao značajnim sa aspekta izgradnje skladišta koje je fokusirano na kupca/korisnika usluga. Sa aspekta CRM-a je vrlo važno primeniti predložena rešenja, jer se samo na taj način može dobiti stvarna slika o kupcima/korisnicima i njihovom ponašanju.

## LITERATURA

- [1] O'Brien J. A., Marakas G. M., Management Information Systems, McGraw-Hill / Irwin, New York, 2006
- [2] Jessup L. M., Valacich J. S., Information Systems Today – Second Edition, Pearson Education Inc., New Jersey, 2006.
- [3] Balaban N., Ristić Ž., Đurković J., Trninić J., Informacioni sistemi u menadžmentu, Savremena administracija, Beograd, 2002.
- [4] Inmon W.H., Building the Data Warehouse – Fourth Edition, Wiley Publishing, Indianapolis, 2005.
- [5] Rojger R. J., Geatz M. W., Data Mining: A Tutorial-based Primer, Addison Wesley, USA, 2003.

- [6] Han J., Kamber M., Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, San Francisco, 2001.
- [7] Todman C., Designing a Data Warehouse: Supporting Customer Relationship Management, Prentice Hall, USA, 2000.
- [8] Ballard C., Herreman D., Schau D., Bell R., Kim E., Valenic A., Data Modeling Techniques for Data Warehousing, IBM Corp., <http://www.redbooks.ibm.com>, 1998.
- [9] Jensen C.S., Clifford J., Gadia S.K., Segev A., Snodgrass R.T., A Consensus Glossary of Temporal Database Concepts, ACM SIGMOD Record, 1992.
- [10] Kimball R., It's Time for Time - Data Warehouse Architect (DBMS online), Miller Freeman, USA, 1997.
- [11] Kimball R., The Data Warehouse Toolkit, John Wiley & Sons, USA, 1996.
- [12] Kimball R., Slowly Changing Dimensions - Data Warehouse Architect (DBMS online), Miller Freeman, USA, 1996.
- [13] Kimball R., Ross M., The Data Warehouse Toolkit (Second Edition)- the Complete Guide to Dimensional Modeling, John Wiley & Sons, USA, 2002.
- [14] Kimball R., Surrogate Keys - Data Warehouse Architect (DBMS online), Miller Freeman, USA, 1998.



mr Dragana Čamilović, Fakultet za menadžment BK, Beograd

Oblast interesovanja: skladišta podataka, data mining, sistemi za podršku odlučivanju

---



---

**info m**

## UPUTSTVO ZA PRIPREMU RADA

Tekst pripremiti kao Word dokument, A4, u kodnom rasporedu 1250 latinica ili 1251 ćirilica, na srpskom jeziku, bez slika.

Naslov, abstrakt i ključne reči dati na srpskom i engleskom jeziku.

Autor(i) treba da obavezno prilože svoju fotografiju, navede instituciju u kojoj radi i oblast kojom se bavi.

Jedino formatiranje teksta je normal, **bold**, *italic*, **bolditalic**, velika i mala slova.

Mesta gde treba ubaciti slike naglasiti u tekstu (Slika 1...)

Proveriti da li su poslate sve slike!

Slike pripremiti odvojeno, VAN teksta, imenovati ih kao u tekstu, u sledećim formatima: vektorske slike - cdr.

(ako ima teksta u okviru slika pretvoriti u krive), ai, fh, eps (šeme i grafikoni), rasterske slike: tif, psd, jpg u rezoluciji 300 dpi 1:1 (fotografije, ekranski prikazi i sl.)

Molimo vas da obratite pažnju na veličinu i izgled slika (prema koncepciji časopisa)

---



---