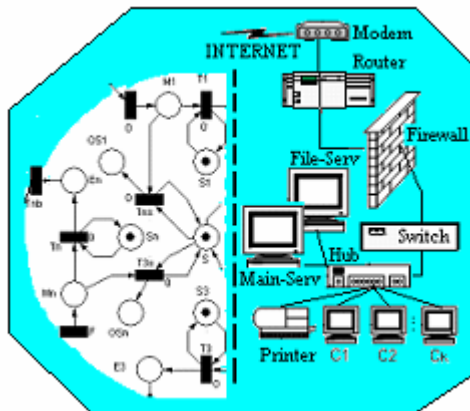


საქართველოს ტექნიკური უნივერსიტეტი

ბია სურგულაძე, ლია კეტრიაშვილი

**მონაცემთა საცავი
ინტერნეტული ბიზნესის
სისტემებისთვის
(DataWarehouse+ OLAP)**



თბილისი 2007

გია სურგულაძე



სტუ-ს ინფორმატიკის ფაკულტეტის სრული პროფესორი, ტექნიკის მეცნიერებათა დოქტორი, ინფორმატიზაციის საერთაშორისო აკადემიის აკადემიკოსი. არის 190 სამეცნიერო-პედაგოგიური ნაშრომის ავტორი (10 მონოგრაფია, 10 სახელმძღვანელო, 20 დამხმარე სახელმძღვანელო /ლექციების კონსპექტი, 25 ელექტრონული წიგნი სტუ-ს ვებ-გვერდზე, 125 სტატია/თეზისი)

ბაზების, ობიექტ-ორიენტირებული დაპროგრამების, მართვის საინფორმაციო სისტემების დაპროექტებისა და პეტრის ქსელების გამოყენებითი თეორიის საკითხებზე. გერმანულ-ქართული ერთობლივი სასწავლო-სამეცნიერო ცენტრის („GeoGer” Center) „ინფოტექნოლოგიები“ დირექტორი (ბერლინის ჰუმბოლდტისა და ნიურნბერგ-ერლანგენის უნივერსიტეტებთან), საქართველოს ინფორმატიზაციის ტექნოლოგთა კავშირის თავმჯდომარე, ყიული შარტავას სახელობის პრემიის ლაურეატი ტექნიკის დარგში.

ლილი პეტრიაშვილი



სტუ-ს ინფორმატიკის ფაკულტეტის „მართვის ავტომატიზებული სისტემების“ კათედრის ასოც-პროფესორი, ტექნიკის მეცნიერებათა კანდიდატი, კათედრის გერმანულენოვანი ჯგუფების კურატორი. არის 20 სამეცნიერო ნაშრომის, 1-მონოგრაფიის, 4 დამხმარე სახელმძღვანელოს და ლექციების კონსპექტის (გერმანულ ენაზე) ავტორი ოპერაციული და საოფისე სისტემების, სტატისტიკური მოდელირების, მასობრივი

მომსახურების სისტემებისა და პეტრის ქსელების თეორიის საკითხებზე. ასპირანტურაში სწავლების პერიოდში 1995-2000 წლებში ოჯახთან ერთად ცხოვრობდა გერმანიაში, სტაჟირებას გადიოდა ნიურნბერგ-ერლანგენის უნივერსიტეტში, „ოპერაციული სისტემებისა და ქსელების კათედრაზე“. პარალელურად დაამთავრა ერლანგენის გერმანული ენის შემსწავლელი ინსტიტუტი, აქვს ატესტატი-სერტიფიკატი. არის ორი შვილის დედა.

სარჩევი

შესავალი	4
თავი I. ბიზნესის მართვის განაწილებული საინფორმაციო სისტემების ანალიზი	8
1.1. ქსელური ინტეგრირებული საინფორმაციო მართვის სისტემების დაპროექტების ამოცანა	8
1.2. ელექტრონული ბიზნესი	10
1.3. ელექტრონული კომერცია	11
1.4. ელექტრონული კომერციის ინფრასტრუქტურული სისტემა	14
1.5. მონაცემთა საცავის ანალიზის OLAP-ნსტრუმენტი	17
1.6. მონაცემთა საცავის შექმნის და მოდელირების ინსტრუმენტი	24
1.7. საცავების მმართველი IBM DB2 Warehouse Manager სისტემა	26
II თავი. მონაცემთა საცავის დაპროექტება	28
2.1. მონაცემთა საცავის სტრუქტურა და ძირითადი ელემენტები	28
2.2. მონაცემთა საცავის დაპროექტების მეთოდურ -ორგანიზაციული საფუძვლები	32
2.3. მონაცემთა ტიპიზაციის, კლასიფიკაციისა და კატალოგიზაციის ამოცანის გადაწყვეტა	40
2.4. მოთხოვნების წინმსწრები ანალიზისა და ტრანზაქციების სინქრონიზაციის სერიალიზაციის ამოცანის გადაწყვეტა	45
2.5. მონაცემთა საცავის პროგრამული კომპონენტები	50
2.6. ოპერატიულ მონაცემთა ბაზის ინფორმაციის კონვერტირების ინსტრუმენტი	52
2.7. OLAP კუბში მონაცემთა აგრეგაცია	55
 ლიტერატურა	 59

შესავალი

კომპიუტერული და ქსელური ინდუსტრიის განვითარებამ ბოლო ათწლეულებში მნიშვნელოვან შედეგებს მიაღწია, რამაც თითქმის მთლიანად შეცვალა მართვის საინფორმაციო სისტემების აგების ტექნოლოგია და ინსტრუმენტული საშუალებანი, გაჩნდა ახალი ცნებები და ტერმინები, რომლებიც თანამედროვე კონცეფციებსა და პროექტებს ედება საფუძვლად.

ორგანიზაციული სისტემების მართვის პრობლემების და ამოცანების გადასაწყვეტად საჭირო ხდება ახალი კომპიუტერული და ინფორმაციული ტექნოლოგიების გამოყენება, რომელთა შორის ერთ-ერთი აქტუალური და მნიშვნელოვანი მიმართულებაა მონაცემთა საცავების (data warehouse) აგება. ბოლო პერიოდში იგი ითვლება, განსაკუთრებით პერსპექტიულ და დინამიკურ მიმართულებად საინფორმაციო სისტემების დაპროექტებაში, იგი უკავშირდება მრავალდონიან განაწილებულ საინფორმაციო სისტემის შექმნას.

მონაცემთა საცავი განიხილება როგორც რომელიმე კონკრეტული ორგანიზაციის ან დიდი საწარმოსთვის განკუთვნილი სპეციალური სუპერ-ბაზა, სადაც მიმდინარე ოპერატიული სამუშაოს შესრულებისას თავს იყრის ქრონოლოგიურ ინფორმაციათა მთელი სპექტრი, რომელთა დანიშნულებაცაა მომხმარებლისთვის ინტერნეტ გვერდებზე მიზნობრივად განლაგებული ტექსტური, გრაფიკული და აუდიო-ვიზუალური საინფორმაციო ბლოკების მიწოდება.

თავისუფალ საბაზრო ეკონომიკის პირობებში ბიზნესის მართვის სისტემებისათვის გადაწყვეტლებათა მიღების ხელშემწყობი ეფექტური მექანიზმების დამუშავება და კვლევა თანამედროვე ქსელური საინფორმაციო ტექნოლოგიების ინტეგრირებული გამოყენებით, ერთ-ერთი მნიშვნელოვანი მიმართულებაა. ელექტრონული ბიზნესისა და კომერციის მომხმარებელს მიეცემა ტექნიკური საშუალება ინტერაქტიულ-დიალოგური კრიტერიუმებთ განსაზღვროს მისთვის საჭირო ინფორმაცია. ამ მიზნით ანალიზური პროცესების სისტემა ეყრდნობა მრავალგანზომილებიან მონაცემთა სტრუქტურებს.

სხვადასხვა საინფორმაციო წყაროებიდან მიღებული მონაცემები ტრანსფორმაციის, კონვერტაციის და ინტეგრირების შემდეგ თავსდება საცავის ობიექტ-ორიენტირებულ მონაცემთა ბაზებში (MS SQL

Server, MS Access, InterBase ან სხვ.), ვინდოუსის(MS_Office, BorlandC++Builder, C#), Web-აპლიკაციების (ASP.NET, ADO.NET, XML) .NET ტექნოლოგიების საფუძველზე.

წინამდებარე ნაშრომში შემოთავაზებულია ავტორთა მიერ საქართველოს ტექნიკურ უნივერსიტეტსა და გერმანიის უნივერსიტეტებში (ბერლინის ჰუმბოლდტისა და ნიურნბერგ-ერლანგენის ინფორმატიკის ინსტიტუტებში) მიღებული თეორიული და ექსპერიმენტული სამუშაოების სამეცნიერო შედეგები, კერძოდ:

- კვლევის შედეგები ელექტრონული ბიზნესისა და კომერციის მართვის საინფორმაციო სისტემების დაპროექტების თანამედროვე მეთოდებისა და ინსტრუმენტების შესახებ, მათი კლასიფიკაციისა და სტრუქტურული ანალიზის საფუძველზე;

- ბიზნესის განაწილებული მართვის სისტემებისთვის მონაცემთა საცავების ობიექტ-ორიენტირებული დაპროექტების და რეალიზაციის მეთოდები და ინსტრუმენტული საშუალებანი საინფორმაციო ობიექტებისა და მეტაინფორმაციის ლოგიკურად მთლიანი ორგანიზებისა და მიზნობრივი დამუშავების კლასთა-ფუნქციების საფუძველზე;

- განაწილებული მართვის საინფორმაციო სისტემების მახასიათებელთა კვლევის შედეგები მასობრივი მომსახურების მოდელის მრავალარხიანი ღია და ჩაკეტილი სისტემებით. ბიზნესისა და კომერციის განაწილებულ ობიექტებზე, კლიენტ-სერვერ არქიტექტურისათვის გამოკვლეულია მოგების ოპტიმიზაციის საკითხები;

- ბიზნესის განაწილებული საინფორმაციო სისტემების რესურსების ეფექტური მართვის დინამიკური მოდელი სტოქასტიკური პეტრის ქსელის გრაფო-ანალიზური ინსტრუმენტით;

- მონაცემებისა და მათი სტრუქტურების კლასტერიზაციის, კონვერტირებისა და აგრეგაციის ალგორითმული სქემები და პროგრამები;

- ბიზნეს-ობიექტების მომხმარებელთა მოთხოვნების წინმსწრები ანალიზისა და მათი დამუშავების სინქრონიზაციის რელაციური სქემები პეტრის ქსელის გრაფული მოდელის საფუძველზე.

ნაშრომის თეორიული და პრაქტიკული შედეგების გამოყენება შესაძლებელია:

– საწარმოებისა და ორგანიზაციების ბიზნესის მართვის სფეროში განაწილებული ინფორმაციული სისტემების ასაგებად და გადაწყვეტილების მიღების ხელშემწყობი მექანიზმების სრულყოფისათვის;

– განაწილებული სისტემების კომპიუტერული და ქსელური რესურსების ეფექტური მართვის ექსპერტული სისტემების ასაგებად;

– მთლიანი ქსელისა და მისი ცალკეული კომპონენტების დატვირთვის, მწარმოებლობის, სისტემის მდგომარეობათა ალბათობების და სხვა მაჩვენებლების გასათვლელად, მოთხოვნათა შემოსვლისა და მათი მომსახურების ინტენსიურობების და ეკონომიკური მაჩვენებლების გათავალისწინებით.

შესავალში განხილულია მართვის საინფორმაციო სისტემების აგების თანამედროვე ტექნოლოგიებისა და ინსტრუმენტულ საშუალებათა დამუშავებისა და სრულყოფის საკითხების აქტუალობა და მნიშვნელობა. ჩამოყალიბებულია ნაშრომის ძირითადი მიზნები, ამოცანები, მათი გადაწყვეტის გზები და საბოლოო შედეგები.

პირველ თავში წარმოდგენილია დიდი საინფორმაციო სისტემების დაპროექტების თანამედროვე ტექნოლოგიების ანალიზი, რომელთა ძირითად სტრუქტურულ კომპონენტსაც მონაცემთა საცავი წარმოადგენს. მისი დამკვიდრება გამოწვეული იყო იმ აუცილებლობით, რომელიც განაპირობა მართვის დიდ და რთულ სისტემებში განაწილებული ინფორმაციული ბაზების ლოგიკურად ერთიან მონაცემთა „საწყობში“ თავმოყრამ. მონაცემთა საცავში ინახება კომპანიების მუშაობის ისტორიის მთელი საინფორმაციო სპექტრი. ისტორიული მონაცემების შენახვა არის ერთ-ერთი აუცილებელი პირობა და იგი მონაცემთა საცავის მთავარი ღირსებაა.

მომხმარებელს საშუალება ეძლევა ინტერნეტ გვერდებზე არსებული ინფორმაციის დახმარებით ეფექტურად და მიზნობრივად იმოქმედოს ბიზნესის სფეროში. საჭირო ხდება ენერჯის, დროის და, რა თქმა უნდა, ფინანსური რესურსების დაზოგვის ღონისძიებების გატარება, რისი უნიკალური საშუალებაცაა ელექტრონული კომერციის (E-Commerce) გამოყენება. როგორც ცნობილია, იგი გამოიყენება ინტერნეტ-ტექნოლოგიის დახმარებით და უზრუნველყოფს კომპიუტერის მონიტორიდან უშუალოდ მსოფლიო ბაზარზე ყიდვა-გაყიდვის წარმოებას. დისერტაციაში აღწერილია

ელექტრონული კომერციის ინფრასტრუქტურული სისტემა. წარმოდგენილია ეკონომიკური მოვლენის განმსაზღვრელი ცვალებადი ფაქტორები. გაანალიზებულია დამოკიდებულება პროდუქციის მიწოდებას, მოთხოვნასა და ფასს შორის თავისუფალ საბაზარო ეკონომიკის პირობებში.

მეორე თავში შემოთავაზებულია განაწილებული მართვის სისტემების მონაცემთა საცავების დაპროექტების ეტაპების, მისი ძირითადი კომპონენტების, ფუნქციებისა და კლასიფიცირებულ მონაცემთა მეტაინფორმაციის ეფექტური ორგანიზაციის, გარდაქმნისა და ძებნის მეთოდებისა და ალგორითმების აღწერა.

მონაცემთა საცავი ორიენტირებულია საგნობრივ სფეროზე და ორგანიზებულია მონაცემთა ოპერატიული ბაზიდან შემოსულ სტრუქტურულად გადამუშავებულ მონაცემთა ქვესიმრავლეების საფუძველზე. ინფორმაციის წყაროს წარმოადგენს სხვადასხვა ორგანიზაციათა დანართები (აპლიკაციები), რომლებიც გამოიყენებს განსხვავებულ პროგრამულ პლატფორმებს და უკავშირდება ოპერატიულ მონაცემთა ბაზას ინტერნეტის საშუალებით (on-line რეჟიმი).

gsurg@gmx.net,
liapetri@Rambler.ru.

თავი I. ბიზნესის მართვის განაწილებული საინფორმაციო სისტემების ანალიზი

1.1. ქსელური ინტეგრირებული საინფორმაციო მართვის სისტემების დააროქტების ამოცანა

თანამედროვე კომპიუტერული ტექნიკისა და ტექნოლოგიების განვითარების პირობებში განაწილებული მართვის ავტომატიზებული სისტემების ერთ-ერთ მთავარ კომპონენტად მონაცემთა საცავებს (Data Warehouse, Repository) მიიჩნევენ [1]. აქ თავმოყრილია საავტომატიზაციო-საკვლევი მართვის ობიექტის სხვადასხვა ტიპის საინფორმაციო ბლოკები: ტექსტები და სუპერტექსტები, ცხრილები და გრაფიკული დიაგრამები, აუდიო და ვიზუალური მასალები და სხვ. ამგვარად, მონაცემთა საცავი აღნიშნული ტიპების ფაილთა მართვის სისტემის რთულ კონგლომერატს წარმოადგენს.

მონაცემთა შეკრება პირველადი წყაროებიდან, მათი გაფილტვრა (მეთოდურად, სემანტიკურად და ტექნიკურად), ტრანსფორმაცია და კონვერტაცია (მონაცემთა წინასწარ განსაზღვრული სტრუქტურების მისაღებად), მეტაინფორმაციის იერარქიულიად ორგანიზება (მონაცემთა კატალოგებისა და არქივების სამართავად), სტანდარტული და გამოყენებითი პროგრამული პაკეტების შექმნა (მონაცემთა ფუნქციურ-მიზნობრივ დასაბუშაველად), ინტერნეტ-გვერდების უახლესი საინფორმაციო ტექნოლოგიებით დაპროექტება (ფართო მომხმარებელთა მარტივი ინტერფეისების ასაგებად).

და ბოლოს, თვით განაწილებული საინფორმაციო სისტემის კომპიუტერული რესურსების ეფექტური მართვის საკითხების კვლევა (ქსელის არხები, კლიენტ-სერვერული პროცესორები, საინფორმაციო ბლოკები, ბაზები და ფაილები, მათი დაცვისა და განახლების მოდულები და ა.შ.) – უმნიშვნელოვანეს ამოცანათა კლასს შეადგენს თანამედროვე მართვის ინტეგრირებული სისტემების ასაგებად.

წინამდებარე წიგნის მიზანი სწორედ ამ საკითხების კვლევა და დაბუშავებაა.

განაწილებული სისტემების დასაპროექტებლად და სარეალიზაციოდ ჩვენ ვიყენებთ დაპროგრამების ობიექტ-ორიენტირებულ და სტრუქტურულ მეთოდებს, უნიფიცირებული მოდელირების ენის UML-ტექნილოგიას [1,2,4].

1.1 ნახაზზე მოცემულია „ადამიანი-მანქანა“ ტიპის ინტეგრირებული საინფორმაციო მართვის სისტემის (ისმს) ზოგადი სტრუქტურული სქემა. მოცემულ კონტექსტში კვლევის საგანს ამ სისტემის კომპონენტთა (საინფორმაციო ბლოკები, პროგრამული პაკეტები, ინტერფეისები და სხვ.) სიმრავლე და მათი ურთიერთკავშირები წარმოადგენს თანამედროვე საკომუნიკაციო და პროგრამული გარდამქმნელების ბაზაზე.

1.2 ნახაზზე მოცემულია ზემოგანხილული ინტეგრირებული საინფორმაციო მართვის სისტემის გაშლილი ზოგადი სქემა. საპრობლემო სფერო (სს), ჩვენს შემთხვევაში ელექტრონული ბიზნესის ან კომერციის კომპლექსური ობიექტია (მაგ., დიდი სავაჭრო ცენტრი).

სისტემის მომხმარებლები (სმ) კლასიფიცირდება მათი ფუნქციური დანიშნულების მიხედვით (მაგ., სისტემის ადმინისტრატორი, მონაცემთა საცავის ადმინისტრატორი, საბოლოო მომხმარებლები - სავაჭრო ცენტრის ხელმძღვანელები და ფუნქციურ ქვედანაყოფთა სპეციალისტები და ა.შ.). I_1 - ინტერფეისია საპრობლემო სფეროს და სისტემის მომხმარებლებს შორის.

საბაზო ინფორმაციულ ტექნოლოგიას (სიბ) წარმოადგენს განაწილებული საინფორმაციო მართვის სისტემის ძირითადი პროგრამული პაკეტები. ესაა ოპერაციული სისტემა (მაგ., მაკროსოფტის Windows 2000/XP და დოტ-NET პლატფორმები) და მაღალი დონის პროგრამული ენები, რომელთაც აქვს ობიექტ-ორიენტირებული მეთოდები და ინსტრუმენტული საშუალებანი (მაგ., C#.NET, Visual Basic.NET, JavaScript.NET, Visual C++.NET, Borland C++ Builder და სხვ.) [5,6,9].

საინფორმაციო ტექნოლოგიები (სტ-ი, შეიძლება განვიხილოთ ფართო სპექტრით. კონკრეტული ინტეგრირებული საინფორმაციო მართვის სისტემის დაპროექტებისას ობიექტ-ორიენტირებული ანალიზის ეტაპზე დაზუსტდება აუცილებელი პროგრამული პაკეტებისა და მათ შორის საკომუნიკაციო ინტერფეისების შედგენილობა. მაგ., დოტ-NET საბაზო პლატფორმის შემთხვევაში ინფორმაციული ტექნოლოგიები შეიძლება იყოს C#.NET, Visual Basic.NET, MsOffice, ASP.NET (Web-აპლიკაციისთვის), ADO.NET. მონაცემთა ბაზების მართვის სისტემების სახით

გამოყენება MsSQL Server-2000, Oracle, InterBase და ა.შ. ყოველ ინფორმაციულ ტექნოლოგიას აქვს საბაზო ტექნოლოგიასთან კომუნიკაციის საინტერფეისო (I_ქ)= ენა [12,17,49].

მონაცემთა საცავებისა და ბაზებიდან ინფორმაციის ამოსაღებად ფართოდაა მიღებული სტანდარტული სტრუქტურირებადი მოთხოვნების ენის ინსტრუმენტის გამოყენება (SQL). სხვადასხვა ინფორმაციული ტექნოლოგიების (მაგ., დაპროგრამების სხვადასხვა ენები, სხვადასხვა ბაზები) გამოყენებისას მონაცემების ან პროგრამული კოდების თავსებადობისათვის შეიძლება CORBA (Common Object Request Broker Architecture) ან DCOM (Distributed Component Object Model) ტექნოლოგიების გამოყენება [19].

დოტ-NET პლატფორმის შემთხვევაში მაიკროსოფტმა ორიენტირებულად გადაწყვიტა და გაამარტივა ეს პროცესი. მან შექმნა შუალედური გარდაქმნის ენა (IL- Intermediate Language). პროგრამები, რომელთა საწყისი კოდები დაწერილია, მაგალითად C#, VB. ან .NET-ის IL-ენაზე, კომპილირდება მანქანურ კოდში. .NET-ის ყველა ენა ფლობს CTS (Common Type System) მონაცემთა შესათანხმებელ საერთო ტიპებს, რათა ენების სტანდარტიზაცია იქნეს მიღწეული [19,22,23]. ამგვარად, ობიექტური კოდები IL-ენის საშუალებით ისე მიიღება, რომ მათში არაა დაფიქსირებული, თუ რომელ ენაზე დაწერილი საწყისი კოდი. დოტ-NET პლატფორმა მდებარეობს Windows-ოპერაციულ სისტემასა და საავტომატიზაციო სისტემის აპლიკაციას შორის (იხ. ნახ.1.3).

1.2. ელექტრონული ბიზნესი

ელექტრონული ბიზნესი ისეთი საქმიანობაა, რომელიც იყენებს ინტერნეტ/ინტრანეტ საინფორმაციო ქსელებს კომპანიის შიგა და გარე კავშირებისათვის, ახორციელებს სამეურნეო, პარტნიორულ და საშუაშაგლო ურთიერთობებს მოგების მიღების მიზნით. იგი მარკეტინგული სისტემაა ინტერნეტული ტექნოლოგიებით.

ელექტრონული ბიზნესი ფირმებისათვის უზრუნველყოფს Web-საიტების შექმნას, კომპანიის ბიზნეს-პროცესების ინტეგრაციას, მათ

კავშირს დამკვეთებთან და მომხმარებლებთან, ბაზრის მასშტაბების გაფართოებას.

ელექტრონული კომერცია ელექტრონული ბიზნესის მთავარი შემადგენელი ნაწილია. იგი ტექნიკური და ორგანიზაციული ფორმების ერთობლიობაა, რომლის საშუალებითაც შესაძლებელია მატერიალური და ფინანსური აქტივების გადაცემა ბიზნესის ერთი სუბიექტიდან მეორეზე. ბიზნეს-ურთიერთობები იყოფა ოთხ კატეგორიად:

- **Business-to-Business (B2B)**- ელექტრონული კომერციაა საწარმოებს შორის, რომელიც მოიცავს ყველა დონის ინფორმაციულ კავშირებს კომპანიებს შორის. ასეთი სისტემები ძირითადად გამოიყენება წარმოების მომარაგებასა და მზა პროდუქციის გასაღებაში;

- **Business-to-Consumer (B2C)**- ელექტრონული კომერციის ვარიანტია, სადაც მყიდველები წარმოადგენენ კერძო პირებს. მაგალითად- ინტერნეტ-მაღაზია, ფასიანი საინფორმაციო სამსახური და ა.შ.;

- **Consumer-to-Consumer (C2C)**- ელექტრონული კომერციაა სხვადასხვა კერძო პირებს შორის, მაგალითად ელექტრონული აუქციონი;

- **Business-to-Government (B2G)**- ელექტრონული კომერციაა საწარმოსა და სახელმწიფო ორგანოებს შორის.

კომპანიის საბაზრო სტრატეგიის კუთხით ელექტრონული ბიზნესი და კომერცია გვთავაზობს ისეთი ფორმების არსებობას ინტერნეტში როგორცაა ელექტრონული სავიზიტო ბარათი, ელექტრონული კატალოგი, ელექტრონული ვაჭრობა (e-trading), ელექტრონული მაღაზიები, (e-shop), ელექტრონული ბუხჰალტერია და ა.შ.

1.3. ელექტრონული კომერცია

ჩვენი ცხოვრების დღევანდელი დინამიკური რიტმი და ტემპი მრავალ საწარმოს და, უპირველეს ყოვლისა, თვით ადამიანს აიძულებს უფრო მოქნილი იყოს გარემოს ცვლილებათა მიმართ. ეკონომიკური საქმიანობის გლობალიზაციის და უნივერსალიზაციის

პროცესი სულ უფრო და უფრო იწევა წინა პლანზე. საჭირო ხდება ენერჯის, დროის და რა თქმა უნდა, ფინანსური რესურსების დაზოგვის ღონისძიებების გატარება, რისი უნიკალური საშუალებაცაა ელექტრონული ბიზნესისა და კომერციის გამოყენება [3].

ელექტრონული კომერცია (E-Commerce), როგორც ცნობილია გამოიყენება ინტერნეტ-ტექნოლოგიის დახმარებით და მომხმარებელს მიეწოდება ინტერნეტ ქსელის საშუალებით [24].

„E“ – electronic (ელექტრონული) ლათინურიდან, ნიშნავს სწრაფს.

„E“ – economical, ინგლისურად ნიშნავს ეკონომიურს.

„E“ – extended business, ნიშნავს, „ბიზნესი საზღვრების გარეშე“.

ამ თვალსაზრისით ელექტრონული კომერცია არის სწრაფი და ბიზნესის ეკონომიური სახე, რომელიც არ ცნობს საზღვრებს. ინტერნეტის შესახებაც არსებობს სხვა-დასხვა მოსაზრებები. [26].

დღეს მსოფლიოში საკმაოდ პოპულარული თემაა ელექტრონული კომერცია – კომპიუტერის მონიტორიდან უშუალოდ მსოფლიო ბაზარზე ყიდვა-გაყიდვის წარმოება. მთელი რიგი კომპანიების საშუალებით წარმატებით იქნა ორგანიზებული თანამედროვე გასაღების ბაზარი. [16] სავაჭრო ორგანიზაციების ინტერნეტში განთავსების მიზეზი რამდენიმეა: ინტერნეტ მაღაზია ფუნქციონირებს მთელი დღე და ღამე (24სთ). მისაწვდომია მსოფლიოს ყველა ქვეყნის მოსახლეობისთვის და უზრუნველყოფს მომხმარებლის ინდივიდუალურ მომსახურებასაც. კორპორაციები IBM, Oracle, Microsoft დიდი ხანია დაინტერესდნენ და პერსპექტივით უყურებენ ამ მიმართულებას, აუძვობენ თავიანთ პროდუქციას საწარმოებლად და ამ სფეროში დაკავებულ ბიზნესმენებს სთავაზობენ სხვადასხვა პროგრამულ უზრუნველყოფას (IBM NET, Commerce, Oracle ICS და სხვა)[25].

კომერციის ელექტრონული სისტემის რეალიზაციისათვის საჭიროა კომპლექსური პროგრამული სისტემა: რამდენიმე პროგრამული ენა და ტექნოლოგია (HTML, Java, JavaScript, JDBC, SQL, XML, მონაცემთა ბაზებისა და ვებ-ინსტრუმენტებით (ASP.NET).

ბრაუზერები (browser) – ვებ კვანძების დათვალიერების პროგრამები მომხმარებელს უზრუნველყოფს ინტერნეტში საძიებო

და სხვადასხვა ვებ-კვანძების და გვერდების დასათვალიერებლად. გრაფიკული ბრაუზერები Netscape და Internet-Explorer საშუალებას იძლევა მომხმარებელმა დაინახოს და „მოისმინოს“ ის ვებ-გვერდები, რომლებიც შეიცავს დინამიკურ ინფორმაციას, აუდიო და ვიდეო ფაილებს.

ვებ-სერვერების უმრავლესობას შეუძლია ერთდროულად დაამუშაოს რამდენიმე მოთხოვნა, ე.ი. ერთ ვებ კვანძს შეუძლია ერთდროულად რამდენიმე მომხმარებელი მიიღოს.

ვებ-სერვერს და მომხმარებელს ესაჭიროება ინფორმაციის გაცვლის ერთიანი პროტოკოლი (Hyper-Text Transfer Protocol HTTP). კლიენტ – სერვერის ურთიერთობის დროს ვებ გვერდის მოთხოვნა იგზავნება HTTP ფორმატში, ვებ სერვერი ინტერპრეტაციას უკეთებს ამ მოთხოვნას და უბრუნებს ბრაუზერს მონაცემებს.

ფირმა Visa International და კომპანია Master Card International პროტოკოლ SET-ში (Secure Electronic Transaction – დაცული ელექტრონული ტრანზაქცია) ხედავს ინტერნეტში ვაჭრობის ძლიერ ინსტრუმენტს, რომელიც აუმჯობესებს ქსელში ჩატარებული ოპერაციების უსაფრთხოებას და აზღვევს გარიგებაში მონაწილე ორივე მხარეს [23].

SET-მომავლის პროტოკოლია, დღეს კი ინტერნეტში ელექტრონული ტრანზაქციების ჩატარება ხდება SSL-ის (Secure Socket Layer) გამოყენებით. SSL არის შიფრაციის პროტოკოლი და პირველ რიგში გამოიყენება ქსელში გადაცემული ინფორმაციის დასაცავად. ელექტრონული კომერციის პრობლემები კი უფრო ფართო სპექტრს შეიცავს, ვიდრე შიფრაცია. აქ შეიძლება საქმე ეხებოდეს ისეთ პრობლემებს, როგორიცაა კლიენტის მიერ შეთანხმების პირობის ან მაღაზიის სინამდვილის შემოწმება. ამიტომ მსოფლიოს უდიდესი კომპანიები დიდ კაპიტალს აბანდებენ ამ მიმართულების გასანვითარებლად და რადიკალურ ფორმებს მიმართავენ მსოფლიო ბაზარზე საკუთარი ადგილის დასამკვიდრებლად, თუმცა მომავალი ციფრული ეკონომიკის ბირთვი (Digital economy) არის არა მხოლოდ სავაჭრო არამედ საქმიანი ტრანზაქციების ჩატარების საშუალება, რაც გულისხმობს ინდივიდუალურ და კორპორაციულ კონტრაქტებს, ინვესტიციების და კაპიტალის გადაადგილებას [17].

1.4. ელექტრონული კომერციის ინფრასტრუქტურული სისტემა

ელ-კომერციის მუშაობა დამოკიდებულია ქსელის გარეგნულ ინფრასტრუქტურაზე. ეს უკანასკნელი მოიცავს ინფორმაციის გადაცემის მატარებლებს [23]. ამიტომ მასში შედის ინტერნეტი, საკაბელო ტელევიზია, ტელეკომუნიკაციური და კერძო კორპორაციული ქსელები.

საწარმოო ინფრასტრუქტურა ფოკუსირდება პროდუქციაზე და იმაზე, რაც საჭიროა მის შესაქმნელად. გაყიდვის ინფრასტრუქტურის მიზანია შეათანხმოს საქონელი და მომსახურება მომხმარებელთან. მომსახურების ინფრასტრუქტურა კი მოიცავს გაყიდვის, გაყიდვის შემდგომი მხარდაჭერის და უსაფრთხოების პროცესებს [24].

ქსელის მუშაობა მთლიანად დამოკიდებულია პროტოკოლებზე – ქსელის მუშაობის წესებზე. პროტოკოლები განსაზღვრავს. თუ როგორ უკავშირდება ელ-კომერციის გამოყენებითი ნაწილი ქსელს, თუ როგორ იყოფა მონაცემები გამოყენებითი პაკეტებად საკაბელო გადაცემებისთვის. ამასთანვე როგორი ელექტრონული სიგნალები წარმოადგენს მონაცემებს ქსელურ კაბელში. მონაცემები დაჯგუფებულია ციფრული ქსელებისათვის გადასაცემად. პაკეტები ამ მონაცემების გარდა შეიცავს მათზე საკონტროლო ინფორმაციასაც. პროტოკოლის შემუშავებისას განისაზღვრება ის, თუ როგორ გაცვლის იგი მონაცემებს მეზობელ დონეებს შორის. მეცნიერულად დასაბუთებული და ეფექტურად ფუნქციონირებადი ელ-კომერციული სისტემა გვთავაზობს ძირითადი ელემენტების ჯგუფს, რომელიც უზრუნველყოფს კონკრეტული პირობების შესრულებას [25].

ელექტრონული კომერციის ძირითადი ელემენტები იყოფა ორ ჯგუფად: პირველში განთავსებული ელემენტები წარმოადგენს ელექტრონული კომერციის ინფრასტრუქტურული სისტემის ყველა კომპონენტს, ხოლო მეორე ჯგუფში განთავსებული ელემენტები კი მათი რეალიზაციის ორგანიზაციული ფორმის მრავალსახეობას.

პირველი ჯგუფის ძირითად ელემენტები მნიშვნელოვნად განსხვავდება იმ ელემენტებისგან, რომელთაც ადგილი ჰქონდა ტრადიციული მაღაზიების შემთხვევაში. ელ-კომერციის სისტემა იყენებს აბსოლიტურად განსხვავებულ ბიზნეს-საშუალებებს [25].

ელექტრონული კომერციის ინფრასტრუქტურული სისტემის ძირითად ელემენტებს აქვს 1.4 ნახაზზე წარმოდგენილი სახე:

- სპეციალური პროგრამული უზრუნველყოფა;
- მონაცემთა ბაზების მართვის სისტემა;
- სატელეკომუნიკაციო კავშირები;
- უსაფრთხოების სისტემა საქონლის ყიდვა-გაყიდვის და მომსახურების სფეროში;
- იურიდიული სამართლიანობის უზრუნველყოფა;
- ვირტუალური საბაზო სისტემა;
- სპეციალური საგადამხდელო სისტემა;
- საწყობის ავტომატური მომსახურება;
- საქონლის მიწოდებისა და მომსახურების სისტემა;
- საფინანსო ინსტიტუტები (საბროკერო და სხვ.);
- საგადასხადო სისტემა და საბაჟო ტარიფები;
- მარკეტინგული მომსახურება, რომელიც მოიცავს: სარეკლამო თუ ფასების რეგულირების და დიზაინის განყოფილებებს: (web-გვერდები, web-სერვერი).

შემოვიტანოთ რამდენიმე კომპონენტის განმარტება.

ელექტრონული მაღაზია. იგი განთავსებულია ინტერნეტ-ქსელში, web-სერვერის სახით. ასეთი სახის მაღაზიის შექმნის მთავარი მიზანია უზრუნველყოს, დროის მინიმალურ შუალედში საქონლის გასაღება და დახმარება გაუწიოს ინტერნეტის სხვა მომხმარებლებს, რათა იაფად და სწრაფად შეარჩიონ მათთვის სასურველი პროდუქცია, ასევე ყველა მომხმარებელს საშუალება აქვს აწარმოოს შეკვეთები.

სპეციალური პროგრამული უზრუნველყოფა. პროგრამული უზრუნველყოფისთვის გამოიყენება დაპროგრამების შემდეგი ენები – Java, HTML, XML და ა.შ. მონაცემთა შეტანა-გამოტანის შაბლონები, დიზაინი და web-გვერდების მომზადების საშუალება, სპეციალური პროგრამული უზრუნველყოფა და ა.შ.

იურიდიული უზრუნველყოფა.

ელექტრონული კომერციის ორგანიზაცია, პირველ რიგში ბაზირებული უნდა იყოს ტრადიციულ იურიდიულ ნორმებით და წესებით, მეორე რიგში კი უნდა მოხდეს ახალი სპეციალიზირებული სამართლის ინსტიტუტების მიერ მიღებული ცვლილებების და

პროცედურების გათვალისწინება. ამას გარდა აქტუალურია კანონმდებლობის უნიფიკაცია, აგრეთვე პროცედურების და წესების გამარტივება, რომლებიც გამოიყენება სხვადასხვა ქვეყნებში. იგი ხელს უწყობს თანამშრომლობას ბიზნესის წარმოებასა და შესაბამის სახელმწიფო მართვის სტრუქტურებს შორის[23].

სპეციალური საგადასახადო სისტემა. ინტერნეტის საშუალებით გადახდის წარმოება ხორციელდება სხვადასხვა ბარათების დახმარებით.[24]

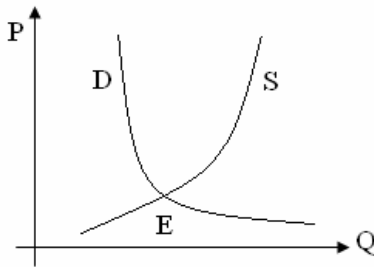
ინტერნეტ ქსელის შესაძლებლობების გამოყენება მომხმარებლის უშუალოდ აქტიური ჩარევით. ეს შესაძლოა იყოს, როგორც მინიმუმი web – გვერდი ან საკუთარი web- სერვერი, ხოლო როგორც მაქსიმუმი – უსაზღვრო შესაძლებლობები, რაც გამოიხატება ინტერაქტიული მაღაზიების არსებობაში და მსხვილი ფირმების ხელმძღვანელებთან უშუალო კავშირებში.

როდესაც ვსაუბრობთ ეკონომიკაზე და მასში მიმდინარე პროცესებზე, აუცილებელია გავითვალისწინოთ რამდენიმე ძირითადი პრინციპი:

ეკონომიკურ მოვლენას განსაზღვრავს მრავალი ცვალებადი ფაქტორი, ამიტომ ამ მოვლენის შესწავლის გასაადვილებლად აფიქსირებენ (უცვლელად თვლიან) ყველა ფაქტორს გარდა ერთისა და სწავლობენ თუ როგორ რეაგირებს მოვლენა ამ ერთი ფაქტორის ცვლილებაზე სხვა ფაქტორების უცვლელობის პირობებში.

მაგალითად, ყველასათვის ცნობილია, რომ მიწოდება, მოთხოვნა და ფასი ურთიერთდაკავშირებულია. იზრდება ფასი კლებულობს მოთხოვნა და პირიქით კლებულობს ფასი მოთხოვნა იზრდება, ეს ურთიერთდაკავშირებული ეკონომიკაში აქსიომად ითვლება (არის გამონაკლისი შემთხვევები, როდესაც ეს პრინციპი ირღვევა).

მოთხოვნის და მიწოდების განმსაზღვრელი ცენტრალური ფაქტორი ფასია. გამყიდველებისთვის უკეთესია მღალი ფასი, ხოლო მყიდველისათვის – დაბალი, თუ ყველა სხვა პირობა ერთნაირია. თუდავადებთ ერთმანეთს გრაფიკებს, მოთხოვნის და მიწოდების მრუდები გადაიკვეთება. 1.5 ნახაზზე ამას შემდეგნაირი სახე აქვს:



ნახ.1.5.

D – მოთხოვნის მრუდი; S – მიწოდების მრუდი; P- ფასების ღირებულება; Q - მიწოდებული საქონლის რაოდენობის ღირებულება; E- წონასწორობის წერტილი.

მოცემული ნახაზიდან[26] ჩანს, რომ არსებობს ერთადერთი ფასი, რომლის დროსაც მომხმარებელთა და მიმწოდებელთა ინტერესები ერთმანეთს ემთხვევა. ამ ფასს წონასწორობის საბაზრო ფასს (Market Equilibrium Price) უწოდებენ. იგი ზუსტად იმ წერტილშია სადაც გადაიკვეთება მოთხოვნისა და მიწოდების მრუდი. ე.ი. წონასწორობის წერტილი – ეს ისეთი წერტილია, რომელშიც მოთხოვნისა და მიწოდების სიდიდეები ერთმანეთის ტოლია, ხოლო საბაზრო ფასი - ეს ისეთი ფასია, რომელიც განისაზღვრება მოთხოვნისა და მიწოდების საშუალებით.

1.5. მონაცემთა სავაჭრო ანალიზის OLAP- ინსტრუმენტი

მონაცემთა მრავალგანზომილებიან კომპლექსური ანალიზის ტექნოლოგიას უწოდებენ **OLAP** (Online Analytical Processing) ტექნოლოგიას, რომელიც ასე განიშარტება: “მონაცემთა ოპერატიული ანალიზი“, მონაცემთა საცავში იგი წარმოადგენს გასაღებურ კომპონენტს.

OLAP – ინსტრუმენტი პირველად მონაცემებს ასახავს ინფორმაციის სახით, რომლის დახმარებითაც შესაძლებელი ხდება საწარმოო მოცულობის შესახებ ვიქონით რეალური წარმოდგენა. ამავე დროს იგი უნიკალური ინსტრუმენტია, რომელიც საშუალებას გვაძლევს სხვადასხვა ჭრილით ჩავატაროთ ინფორმაციის მრავალგანზომილებიანი ანალიზი.

1.5.1. OLAP მრავალგანზომილებიანი წარმოდგენა = კუბი.

OLAP წარმოადგენს სწრაფმოქმედ და ხელსაყრელ საშუალებას საქმიანი ინფორმაციის დათვალიერებისა და ანალიზისათვის [57,58]. იგი არ არის ცალკე აღებული პროგრამული პროდუქტი, პროგრამირების ენა ან რომელიმე კონკრეტული ტექნოლოგია.

OLAP–ს გააჩნია საკუთარი კონცეფცია, პრინციპები და მოთხოვნები, რომელიც ეყრდნობა პროგრამულ პროდუქტს და ანალიტიკოსებს უადვილებს მონაცემებთან ურთიერთობას.

ანალიტიკოსი არის განსაკუთრებული მომხმარებელი კორპორაციული ინფორმაციით. ანალიტიკოსის ამოცანაა დაამყაროს წესრიგი დიდ მონაცემთა მასივში, ამიტომ მისთვის მნიშვნელოვანი არ არის ცალკეული ფაქტები. მონაცემთა საცავში არსებული ცალკეული ფაქტებით შეიძლება დაინტერესდეს მაგალითად, ბუხჰალტერი ან რომელიმე განყოფილების ხელმძღვანელი, რომელთა კომპეტენციაშიც შედის საქმიანი შეთანხმებები.

ანალიტიკოსისათვის კი ერთი რომელიმე სახის მონაცემი თითქმის არაფერს წარმოადგენს, მისი ინტერესის სფერო შეიძლება გახდეს დროის რალაც შუალედში (მაგ., ერთი თვის, ერთი წლის) რომელიმე მსხვილი ობიექტის, ყველა სახის საქმიანი შეთანხმება.

ამავე დროს ანალიტიკოსს შეუძლია არ გაამახვილოს ყურადღება ისეთ ინფორმაციაზე, როგორცაა მაგ., მომხმარებლის მისამართი, ტელეფონის ნომერი, კონტრაქტის ინდექსი და ა.შ.

მისთვის საჭირო ინფორმაცია წარმოდგენილი უნდა იყოს რიცხვითი მნიშვნელობებით, ეს განაპირობებს მისი მოქმედების ძირითად არსს. ეს უკანასკნელი მიუთითებს იმაზე, რომ ანალიტიკოსი მუშაობს ასეთი სახის ცხრილთან:(ცხ.1.1) ქვეყანა, საქონელი და წელი წარმოადგენს ატრიბუტებს, ხოლო ყიდვა გაყიდვის მოცულობა გამოსახულია რიცხვითი მნიშვნელობით. ანალიტიკოსის ამოცანა მდგომარეობს ატრიბუტებსა და რიცხვით პარამეტრებს შორის დამოკიდებულების გარკვევაში. თუ ცხრილს დავაკვირდებითშევაძინებთ, რომ ცხრილის წარმოდგენა შესაძლებელია სამ განზომილებაში: ერთი ლერძის გასწვრივ დალაგდება ქვეყნები, მეორეზე, საქონელი, ხოლო მესამეზე წელი.

თავისი მნიშვნელობით იგი გამოსახავს სამგანზომილებიან მასივს შესაბამისი ყიდვა-გაყიდვის მოცულობით.

ცხრ.1.1

დონანა	სამრეალო	წელი	მიღვა- გამოღვის შრიფტოვა
ტეშმანია	საყოფაცხო. ელექ.	2001	234
ვენესუელა	საყოფაცხო. ელექ.	2001	456
ატენტინა	საყოფაცხო. ელექ.	2002	321
ტეშმანია	საყოფაცხო. ელექ.	2003	342
ვენესუელა	ტეზინის ნაკეთ.	2003	123
ატენტინა	საყოფაცხო. ელექ.	2001	657
ვენესუელა	ტეზინის ნაკეთ.	2002	143
ტეშმანია	ტეზინის ნაკეთ.	2002	345
ვენესუელა	ტეზინის ნაკეთ.	2001	56
ატენტინა	საყოფაცხო. ელექ.	2003	234
ტეშმანია	ტეზინის ნაკეთ.	2003	309
ატენტინა	ტეზინის ნაკეთ.	2002	345
ვენესუელა	საყოფაცხო. ელექ.	2003	117
ატენტინა	ტეზინის ნაკეთ.	2001	57
ტეშმანია	ტეზინის ნაკეთ.	2002	152
ვენესუელა	საყოფაცხო. ელექ.	2002	672
ატენტინა	ტეზინის ნაკეთ.	2003	212

1.6 ნახაზზე განსხვავებული ფერი გამოხატავს სეგმენტს სადაც იმ წელს არ არსებობს მონაცემი მაგ. გერმანიას 2001

წელს რეზინის ნაკეთობებზე არა აქვს ინფორმაცია. სწორედ ასეთი სახის სამ განზომილებიან მასივს უწოდებენ OLAP – კუბს.

აუცილებლობას არ წარმოადგენს კუბის ყველა ელემენტის შევსება, მაგ., თუ არ იქნება ინფორმაცია, როგორც აღვნიშნეთ გერმანიაში 2001 წელს რეზინის ნაკეთობებზე, მაშინ შესაბამისი უჯრედი უბრალოდ არ იქნება განსაზღვრული OLAP – დანართში და ამავე დროს საჭიროებას არ წარმოადგენს დანართში მოხვედეს მრავალგანზომილებიანი სტრუქტურის ყველა ელემენტი, მთავარია მომხმარებელმა მიიღოს მხოლოდ საჭირო მონაცემები. კუბში მრავალგანზომილებიან მონაცემთა კომპაქტური შენახვა განაპირობებს, რომ არ მოხდეს მენსიერების უსარგებლო დაკარგვა არასასურველ ინფორმაციაზე.

თუ ნამდვილ კუბს, შევადარებთ OLAP – კუბს, მაშინ მათ შორის შევნიშნავთ განსხვავებას – ნამდვილი კუბის ელემენტების რაოდენობა ყველა განზომილებაში ერთნაირია, ხოლო OLAP – კუბი არ არის ზუსტად განსაზღვრული. იგი შეიძლება იყოს როგორც ორგანოზომილებიანი, ასევე სამ და მრავალგანზომილებიანი, დამოკიდებულია თუ როგორი სახის ამოცანასთან გვაქვს საქმე.

OLAP – კუბის თითოეული განზომილება შედგება ეგრეთ წოდებული შრეების ან ცალკეული ნაწილებისგან მაგ., განზომილება „ქვეყანა“ შედგება შრეებისგან: გერმანია, ვენესუელა, არგენტინა და ა.შ. თვითონ მრავალგანზომილებიანი კუბის გამოყენება ანალიზისათვის შეუძლებელია. მაგ., თუ მოცემული გვაქვს ექვს ან ცხრამეტგანზომილებიანი კუბი მათი აღეკვალური გარდაქმნა სირთულეებთან არის დაკავშირებული, ამიტომ გამოყენების წინ მრავალგანზომილებიანი კუბიდან გამოყოფენ, ორგანოზომილებიან ცხრილს ამ ოპერაციას უწოდებენ კუბის „გაკვეთას“. ანალიტიკოსები მრავალგანზომილებიან კუბიდან „ამოკვეთენ“ მათთვის საინტერესო შრეებს. შესაბამისად „გაუკვეთავი“ რჩება ორი განზომილება. როდესაც კუბი შეიცავს რამდენიმე სახის რიცხვით მნიშვნელობას, ერთი სახის მნიშვნელობები განთავსდება ცხრილის ერთ განზომილებაში.

თუ ზემოთ მოცემულ ცხრილს კარგად დავაკვირდებით, შევამჩნევთ რომ იქ მოცემული მონაცემები პირველადი კი არ არის, არამედ შედარებით მცირე ელემენტები დაჯგუფებულია და

წარმოდგენილია სრულყოფილი სახით. მაგ., წელიწადი შეგვიძლია დავყოთ კვარტალებად, კვარტალები თვეებად, თვეები კვირებად, კვირები დღეებად და ა.შ. ქვეყანა შედგება რეგიონებისგან რეგიონები – დასახლებული პუნქტებისგან.

OLAP – ში ასეთი სახის მრავალდონიან გაერთიანებას უწოდებენ იერარქიას. ცალკეული ელემენტი შეიცავს რამოდენიმე სახის იერარქიას: მაგ. დღე-კვირა-თვე ან დღე-დეკადა-კვარტალი. მონაცემთა დაჯგუფება იწყება ყველაზე დაბალი დონიდან და ბოლოს ჯგუფდება მაღალი დონის მონაცემამდე. იმისათვის, რომ დაჩქარდეს ერთი მდგომარეობიდან მეორეში გადასვლის პროცესი და მონაცემთა სხვადასხვა დონეზე დაჯგუფება, ყველა მონაცემი უნდა ინახებოდეს კუბში. ერთი შეხედვით მომხმარებელი ხედავს მხოლოდ ერთ კუბს, მაგრამ შეგვიძლია ვთქვათ, რომ ეს კუბი შეიცავს პრიმიტიული კუბების მთელ სიმრავლეს.

1.5.2. იერარქია ანუ მონაცემთა აგრეგატული სახით წარმოდგენა

ერთ-ერთი მნიშვნელოვანი გარემოება, რამაც **OLAP** – ტექნოლოგია წარმოაჩინა, როგორც ეფექტური და მაღალმწარმოებლური ინსტრუმენტი, ეს არის იერარქიული გაერთიანებების დამსახურება.

ანალიტიკოსი, დამოუკიდებლად ან პროგრამისტის დახმარებით დასვავენ შესაბამის **SQL** – მოთხოვნას და შედეგად ღებულობს მონაცემებს ელექტრონული ცხრილის სახით. ამ დროს წარმოიქმნება დაბრკოლებათა მთელი სიმრავლე, საიდანაც შეიძლება გამოვყოთ რამდენიმე მათგანი:

1. ანალიტიკოსი გარკვეული დროის განმავლობაში ელოდება პროგრამისტისგან ინფორმაციის მიღებას;

2. ინფორმაციის მიწოდება ხდება მხოლოდ ერთი ცხრილის საშუალებით, რის გამოც შეუძლებელია ჩატარდეს სრულყოფილი კვლევა. ამიტომ ერთსა და იმავე პროდუქტზე უნდა ჩატარდეს რამდენიმე გამოკვლევა;

3. როგორც ზემოთ აღვნიშნეთ, ანალიტიკოსს არ სჭირდება ყველა უმნიშვნელო ინფორმაცია, მას სჭირდება მხოლოდ მნიშვნელოვანი ინფორმაცია ერთდროულად. ეს იმას ნიშნავს, რომ კორპორაციულ რელაციური მონაცემთა ბაზების მართვის სისტემის

სერვერმა, რომელსაც ანალიტიკოსი მიმართავს, უნდა მიაწოდოს ყველაზე აქტუალური ინფორმაცია და, ამავე დროს, დაბლოკოს სხვა დანარჩენი ტრანზაქციები.

არსებული პრობლემის გადასაჭრელად OLAP – ტექნოლოგია არის ერთ-ერთი საშუალება. OLAP – კუბი წარმოადგენს მეტა-ანგარიშთა ნაკრებს. გაკვეთილი მეტა-ანგარიში (ანუ კუბი) ანალიტიკოსს საშუალებას აძლევს სხვადასხვა განზომილებაში წარმოადგინოს მისთვის საინტერესო ორგანიზაციის სტრუქტურული მონაცემები ან ანგარიში.

სისტემის სრულყოფილი ანალიზისთვის OLAP – ინსტრუმენტის გამოყენებისას მნიშვნელოვანია 5 ძირითად პრინციპი, მას უწოდებენ FASMI – ტესტს. აქ კონკრეტულად განსაზღვრულია OLAP – პროდუქტის მოთხოვნები.

FASMI – აბრევიატურა განისაზღვრება ტესტის შემადგენელი თითოეული პუნქტისგან:

- Fast (სწრაფი) – OLAP დანართმა უნდა უზრუნველყოს ანალიზური მონაცემების მიღება მინიმალურ დროში. ანალიზი უნდა ჩატარდეს რაც შეიძლება სწრაფად და უნდა მოხდეს ყველა ინფორმაციული ასპექტის გათვალისწინება. ამ პერიოდის ხანგრძლივობა საშუალოდ უნდა იყოს ხუთი წამი და ურფო ნაკლები;

- Analysis (ანალიზი) – OLAP დანართმა მომხმარებელს უნდა მისცეს საშუალება განახორციელოს რიცხვითი და სტატისტიკური ანალიზი;

- Shared (ერთდროულად გამოყენება) – OLAP დანართმა უნდა უზრუნველყოს ერთსა და იმავე დროს რამდენიმე მომხმარებლის ერთდროული მუშაობა. ამავე დროს უნდა შემოწმდეს, რამდენად დაცულია კონფიდენციალური ინფორმაცია;

- Multidimensional (მრავალგანზომილება) – ეს არის OLAP – ის დამახასიათებელი არსებითი თვისება;

- Information (ინფორმაცია) – OLAP დანართმა მომხმარებელს უნდა მისცეს საშუალება, მიიღოს საჭირო ინფორმაცია, დამოუკიდებლად მისი მოცულობისა და შენახვის ადგილისა.

1.5.3. მონაცემთა მრავალგანზომილებიანი შენახვის ტექნიკური ასპექტები.

მრავალგანზომილებიან მონაცემთა საცავში თავმოყრილი აგრეგატული მონაცემები სხვადასხვა დონეზე ერთმანეთისგან განსხვავდება. მაგ., ყიდვა-გაყიდვის მოცულობა ერთ დღეში, თვეში, წელიწადში, საქონლის კატეგორია და ა.შ.

აგრეგატულ მონაცემთა შენახვის მიზანია მოთხოვნათა დაკმაყოფილებაზე დახარჯული დროის შემცირება, რამდენადაც უმეტეს შემთხვევაში ანალიზის და პროგნოზის სფეროს წარმოადგენს არა დეტალური ინფორმაცია, არამედ გაერთიანებული, საბოლოო სახემდე მიყვანილი დაჯგუფებული მონაცემები. ამიტომ მრავალგანზომილებიან მონაცემთა ბაზის შექმნის დროს, ყოველთვის გასათვალისწინებელია, რომ შეტანილი იქნეს რამდენიმე აგრეგატული მონაცემი. გასათვალისწინებელია ის ფაქტიც, რომ ყველა მონაცემის აგრეგატული სახით შენახვა გაუმართლებელია, რადგან მონაცემთა მოცულობაში ახალი განზომილების დამატების შემთხვევაში კუბი სადაც ამ მონაცემის ჩამატება მოხდება, დაიწყებს ზრდას ექსპონენციალურად ამ პროცესს უწოდებენ მონაცემთა მოცულობის ‘ფეთქებად ზომას’, ამიტომ აგრეგატულ მონაცემთა მოცულობის ზომის ხარისხი დამოკიდებულია კუბების რაოდენობაზე და იერარქიის სხვადასხვა დონეზე მიმდინარე გაზომვებზე. ‘ფეთქებად ზომის’ პრობლემის გადასაჭრელად გამოიყენება სხვადასხვა სახის მონაცემთა შენახვის სქემები : MOLAP, ROLAP, JOLAP და HOLAP.

MOLAP – (მრავალგანზომილებიანი **OLAP**) პროდუქტს საფუძვლად უდევს მონაცემთა არარელაციური სტრუქტურა, რომელიც უზრუნველყოფს მონაცემთა მრავალგანზომილებიან შენახვას.

ROLAP – (რელაციური **OLAP**) ასეთ ინსტრუმენტში მრავალგანზომილებიანი სტრუქტურა რეალიზდება რელაციური ცხრილებით, ხოლო მონაცემები ანალიზის პროცესში შესაბამისად შეირჩევა მონაცემთა რელაციური ბაზებიდან ანალიზური ინსტრუმენტით.

JOLAP – ინსტრუმენტის გამოყენებისას თითოეული მოთხოვნის მოდელირება შესაძლებელია, როგორც ერთიან ობიექტთა ჯგუფი, რის განხორციელებაშიც ეხმარება გრაფიკულ ინტერფეისში ოპერაციათა

თანამიმდევრული სტრუქტურა და მოთხოვნათა მოდიფიკაციის განსაზღვრა. **JOLAP** (Java Specification Request-69) გამოიყენება სხვადასხვა სპეციფიკაციაში მაგ., როგორცაა: ჯგუფი **OMG** მეტამოდელისთვის **Common Warehouse Metamodel (CWM)**, ობიექტური მოდელი **Meta Object Facility (MOF)**, **XML Metadata Interchange (XMI)**, და აგრეთვე **Java Metadata Interface (JMI,JSR-40)**. ამ ინტერფეისის არქიტექტურა ერთდროულად **J2EE** და **J2SE** ;

მრავალგანზომილებიანი **OLAP** უზრუნველყოფს მაღალ მწარმოებლურობას, მაგრამ სტრუქტურები არ შეიძლება გამოვიყენოთ დიდი მოცულობის მონაცემთა დამუშავებისთვის, რადგანაც დიდი გაბარიტები მოითხოვს დიდ აპარატულ რესურსებს. ამასთან ერთად ჰიპერკუბის დატვირთვა შეიძლება იყოს ძალიან მაღალი, რაც შემდგომში თავს იჩენს ინფორმაციის მიწოდების სისწრაფეზე.

შეიძლება ვთქვათ პირიქითაც, რელაციური **OLAP** უზრუნველყოფს შენახული მონაცემების დიდი მასივების დამუშავებას, ასე რომ შეიძლება უზრუნველვყოთ უფრო ეკონომიური შენახვა, მაგრამ ამასთან ერთად, მნიშვნელოვან როლს ითამაშებს მრავალგანზომილებიან სამუშაოს შესრულების სიჩქარეზე. მსგავსს მსჯელობას მიეყვართ ახალი კლასის გამოყოფაზე, რომელიც არის **HOLAP** ანალიზური ინსტრუმენტი.

ეს არის ჰიბრიდული მონაცემთა ოპერატიულ-ანალიზური დამუშავების ინსტრუმენტი. ასეთი კლასის ინსტრუმენტი საშუალებას იძლევა შეთანხმდეს ორივე მიდგომა – რელაციური და მრავალგანზომილებიანი. ხელმისაწვდომი გახდება როგორც მრავალგანზომილებიანი მონაცემთა ბაზები ასევე რელაციური მონაცემები.

1.6. მონაცემთა საცავის შიქონის და მოდელირების ინსტრუმენტი

მონაცემთა საცავების შექმნისა და მოდელირების ინსტრუმენტი შეიძლება გამოყენებულ იქნას სხვადასხვა წყაროებიდან მიღებულ მონაცემთა მოგროვებისთვისა ამ ინფორმაციის გარდაქმნისა და აგრეთვე მის გლობალურ ან ლოკალურ დონეზე საცავში ჩასატვირთად.

საცავის სამრეწველო რეჟიმში მართვისათვის გამოიყენება IBM DB2 Warehouse Manager, ინსტრუმენტი, რომელიც წარმოადგენს საცავის ძირითად მექანიზმს[4].

მონაცემთა საცავის დაპროექტების ეტაპზე Data Warehouse Center – წარმოადგენს სათადარიგო ინსტრუმენტულ საშუალებას მონაცემთა საცავის შექმნისათვის. მის შემადგენლობაში შედის Warehouse Schema Modeler და Process Modeler. ისინი გამოიყენება დაპროექტებისთვის, სხვადასხვა სქემათა გენერაციისათვის და მონაცემთა ჩატვირთვისათვის, აგრეთვე შესაძლებელია ჩატარებულ მოქმედებათა გრაფიკული აღწერა, რომელიც სრულდება საცავის აგებისას. Process Modeler ახორციელებს მონაცემთა უპირობო დამუშავებას, არსებულ მოვლენებზე შეტყობინებას და მართვის კალენდარულ დაგეგმვას.

IBM DB2 UDB Data Warehouse Editions – ახალი, მრავალფუნქციური პაკეტია მონაცემთა საცავების ორგანიზებისათვის, რომლის ბაზასაც IBM DB2 Universal Database Version 8.1 Enterprise Server Edition (ESE)–წარმოადგენს, ამავე დროს იგი არის დამატებითი საშუალება საცავის ადმინისტრირებისათვის, მონაცემთა ინტეგრაციისათვის, მართვისა და ანალიზისთვის.

DB2 Data Warehouse Enterprise Edition –პაკეტი ორიენტირებულია ნებისმიერი ზომის მონაცემთა საცავის შექმნაზე ნებისმიერ დარგში, რაც საშუალებას გვაძლევს ავაგოთ მძლავრი საცავი ფართო ანალიზური დანართებისთვის. პაკეტის მოცულობა შეიძლება იყოს 10 ტერაბაიტზე მეტი.

DB2 Data Warehouse Enterprise Edition-ში შეიძლება გამოყენებული იქნას შემდეგ პლატფორმებზე: IBM AIX, Microsoft Windows 2000/2003, Linux Sun Solaris.

IBM DB2 Cube Views –წარმოადგენს გადაწყვეტილების მიღების თანამედროვე საშუალებას, რომელიც OLAP – ოპერაციებს ახორციელებს მონაცემთა ბაზების მართვის IBM DB2 სისტემებში. მისი დახმარებით მომხმარებელს შეუძლია შექმნას მონაცემთა სწრაფად გარდამქმნელი და იოლად მართვადი OLAP – გადაწყვეტილება, რომელიც ხელს უწყობს ამაღლდეს ანალიზური დანართების მთელი სპექტრის წარმადობა.

IBM DB2 Cube Views–გამჭვირვალედ წარმოადგენს მრავალგანზომილებიან მონაცემთა ბაზების სტრუქტურას შემდეგი ფაქტორების დახმარებით:

ქმნის მეტამონაცემებს სხვადასხვა განზომილების, იერარქიის, ატრიბუტების და ანალიზური ფუნქციისათვის;

ხელს უწყობს და ამაღლებს OLAP – ინსტრუმენტის ეფექტურობას;

- DB2–კატალოგში შეტანილი OLAP მეტამონაცემები ამაღლებს ამ ინსტრუმენტების მაღალეფექტურობას;

- DB2–მაღალეფექტური ტექნოლოგიის გამოყენება იმდენად მარტივია, რომ იგი შეიძლება შევადაროთ ჯამურ ცხრილებსა და ანალიზურ ფუნქციებს.

IBM DB2 Query Patroller – მოთხოვნათა მართვის მაღალმწარმოებლური სისტემა, რომელიც ახორციელებს დინამიკურ კონტროლს მოთხოვნათა ნაკადზე და DB2–მონაცემთა ბაზაზე შემდეგი საშუალებებით:

- მოთხოვნათა შორის რაციონალურად უნდა განაწილდეს სისტემის რესურსები, მიუხედავად იმისა როგორი სირთულითაა მოთხოვნა;

- მოთხოვნათა ავტომატური გადაყვანა ლოდინის რეჟიმში;

- კონტრილიდან გამოსულ მოთხოვნათა სისტემიდან გამოყვანა.

Query Patroller - უზრუნველყოფს მონაცემთა საცავე დატვირთვის რეგულირებას, მცირე და დიდ მოთხოვნათა სწრაფად დასაკმაყოფილებლად სისტემის რესურსების ეფექტურ გამოყენებას. აგრეთვე იგი ხელს უწყობს მონაცემთა მოგროვებას და ინფორმაციის ანალიზს მოთხოვნაზე პასუხის გასაცემად.

1.7. საცავეების მმართველი IBM DB2 Warehouse Manager – სისტემა

IBM DB2 Warehouse Manager – უზრუნველყოფს მონაცემთა საცავეების ავტომატიზებულ ორგანიზებას [56]. მისი მიზანია ბიზნესთან დაკავშირებული ინფორმაციის ღრმად ჩაწვდომა და მართვა, რათა სწორად მოახდინოს ბიზნესის წარმოება. ამ ინფრასტრუქტურაში შეიძლება გაერთიანდეს სხვა ინსტრუმენტები და Business Intelligence (ინტელექტუალური ბიზნესი) დანართები,

რომლებიც გვაძლევს იმის გარანტიას, რომ მოვიღებთ იმ დროისთვის არსებული ყველაზე აქტუალური ინფორმაცია, რაც დაგვეხმარება ბიზნესის სხვადასხვა სფეროში გააზრებული გადაწყვეტილების მიღებაში. IBM DB2 – არის მონაცემთა ბაზების მართვის საუკეთესო სისტემა, რომელიც ხელს უწყობს საინფორმაციო საცავის შექმნას.

DB2- საცავის მენეჯერი: გეთავაზობს გაფართოებულ ფუნქციონალურ გაერთიანებას, სადაც საბაზისო DB2-ს შესაძლებლობას ემატება (ETL – Extraction, Transformation and Loading) მონაცემთა გარდაქმნის და ჩატვირთვის მექანიზმები;

- მეტამონაცემთა და საინფორმაციო კატალოგის მართვა. საინფორმაციო კატალოგი იძლევა მეტამონაცემთა დასმული მოთხოვნის გაერთიანების საშუალებას;

- განაწილებულ სისტემებში მონაცემთა ჩატვირთვის და გარდაქმნას- ETL;

მასში გაერთიანებულია QMF for Windows – დანართი, რომელიც ხელს უწყობს DB2 პაკეტისთვის, Windows ან Web ინტერფეისის დახმარებით მოთხოვნების დაყენებას.

აღსანიშნავია აგრეთვე DB2 Warehouse Manager V8-ის შემდეგი მხარე:

- გამოყენებისას იგი საკმაოდ მარტივია, გამოირჩევა მაღალი მწარმოებლურობით;

- ერთდროულად ახორციელებს რამდენიმე ოპერაციას;

- ახდენს საწყის მონაცემთა თავმოყრას და წარმოადგენს გაერთიანებული საცავის სახით;

- ახდენს მონაცემთა სწრაფად შევსებას;

- საინფორმაციო კატალოგისთვის ახალი ინტერფეისის სრულყოფა;

- უზრუნველყოფს IBM AIX და Linux – ოპერაციული სისტემისთვის მონაცემთა საცავის სერვერის დაკავშირებას.

II თავი. მონაცემთა საცავის დაპროექტება

2.1. მონაცემთა საცავის სტრუქტურა და ძირითადი ელემენტები

მონაცემთა საცავის (Data Warehouse - DWH) იდეა, რომელიც წარმოადგენს მართვის საინფორმაციო სიტემების დაპროექტების და რაეალიზაციის ერთ-ერთ უახლეს ტექნოლოგიას, აქტუალური გახდა 90-იანი წლებიდან. ამ იდეის ფუძემდებლად მოიხსენიება ამერიკელი მეცნიერი ვ. ჰ. ინმონი [27].

მონაცემთა საცავი აღიწერებოდა, როგორც „მონაცემთა სუპერმარკეტი“, „სუპერ მონაცემთა ბაზა“. ამ მიმართულებით პირველი პროექტი „ევროპის ბიზნეს ინფორმაციული სისტემა“ 1988 წელს IBM ფირმის მიერ განხორციელდა. გლობალურ მონაცემთა შენახვა და მათი ანალიზი სავაჭრო კომპანიების მუშაობისას არა მარტო ამაღლებს მუშაობის ეფექტურობას, არამედ უძლებს მკაცრ კონკურენციას, რომელიც საერთაშორისო ბაზარზე ყოველწლიურად ძლიერდება.

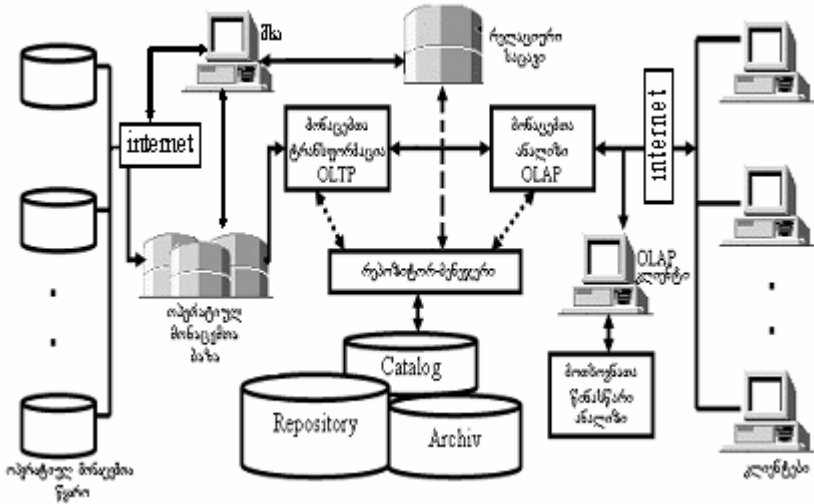
მონაცემთა საცავის ფართო გამოყენება, მსოფლიო ბიზნეს გაერთიანებაში ადასტურებს იმ ფაქტს, რომ ამ ტექნოლოგიაზე დაყრდნობით მოგება ყოველწლიურად იზრდება ათობით მილიარდი დოლარით. ბანკებში ჩატარებულმა ანალიზმა აჩვენა, რომ საბაზო ტრანზაქციების და სასურველ მონაცემებზე არსებული ინფორმაციის მოცულობა ძალიან დიდია. პირველადი ანალიზისათვის უნდა მომზადდეს გაფართოებული მონაცემები და მიეცეს ინდექსაცია. ამისათვის საჭიროა ფართო კომპიუტერული რესურსები, რომელიც საშუალებას იძლევა მცირე დროის განმავლობაში შესრულდეს ძებნა რამდენიმე ცხრილიში, რომელიც ათობით მილიონ ჩანაწერს შეიცავს და განხორციელდეს მონაცემთა შერჩევა.

მონაცემთა საცავი არის კომპლექსური სისტემა, რომელიც შედგება შემდეგი ძირითადი ფუნქციური ბლოკებისაგან:

- მონაცემთა განაწილებული, რელაციური ბაზების მართვის სისტემები;
- ინფორმაციის წყაროებიდან ოპერატიულ მონაცემთა ჩატვირთვის და გარდაქმნის საშუალება;

- საცავის დაპროექტების მეთოდური და ინსტრუმენტული საშუალებანი;
- საცავის აგებისა და მოდიფიკაციის საშუალებანი;
- საცავის მეტამონაცემთა იერარქიული ორგანიზების ჰიპერლინკური საშუალებანი;
- საცავის ფუნქციურ მომხმარებელთა მოთხოვნების წინასწარი ანალიზისა და ტრანზაქციების ეფექტურად შესრულების დაგეგმვის საშუალებანი;
- საცავის საინფორმაციო ბლოკებისა და არქივის ოპერატიული ანალიზის ინსტრუმენტული საშუალებანი.

2.1 ნახაზზე მოცემული გვაქვს განაწილებული ავტომატიზებული მართვის სისტემის მონაცემთა საცავის ზოგადი სქემა.



ნახ.2.1

მონაცემთა საცავის მუშაობის პრინციპი ასეთია: პირველ ეტაპზე DWH-ს ტექნოლოგიის გამოყენების საშუალებით, რელაციურ ბაზებში ერთად თავმოყრილი მონაცემები ლაგდება გარკვეული სტრუქტურული თანამიმდევრობით, ხდება მათი „დაწმენდა“. მეორე ეტაპზე წარმოებს მათი ტექნოლოგიური დამუშავება მონაცემთა ოპერატიული ანალიზის OLAP – ტექნოლოგიის გამოყენებით. ხოლო მესამე ეტაპზე ეს მონაცემები მომხმარებლებს მიეწოდებათ ინტერნეტის საშუალებით.

ინფორმაციული ბლოკები, რომლებიც მონაცემთა საცავებშია განაწილებული, მიზანმიმართულად თავსდება ინტერნეტ – გვერდებზე და ხელმისაწვდომია ფართო მომხმარებლისთვის.

მონაცემთა საცავის კონცეფციის ავტორი, ვ. ინმონი აღნიშნავდა, რომ მონაცემთა საცავი არის: „საგნობრივ ორიენტირებულ მონაცემთა ქრონოლოგიური ნაკადი, რომლის მიზანსაც ორგანიზებული მართვა წარმოადგენს“ [28,29].

მონაცემთა რელაციურ მოდელებზე აგებულ ბაზებისა და საცავებისათვის SQL (Structured Query Language) მონაცემთა სტრუქტურირებადი ენა არის ერთ–ერთი ეფექტური, საუკეთესო საშუალებაა მონაცემთა მანიპულირებისათვის. იგი აღიარებულია საერთაშორისო სტანდარტად, ამიტომაც ინფორმაციული საცავებისათვის მიზანშეწონილია რელაციური პლატფორმა.

რელაციური ცხრილებით წარმოდგენილ მონაცემებსა და საინფორმაციო ბლოკებს შორის არსებობს ლოგიკურ ურთიერთმიმართებათა ოთხი სახე, რომლებიც გამოიყენება კონცეპტუალური მოდელირებისათვის:

1. ერთი-ერთთან (1:1) – განხილულ ატრიბუტს აქვს იერარქიული ატრიბუტის ჩვეულებრივი რეკვიზიტის სახე. მაგ., ორგანიზაციის სრული დასახელება, რომელიც კავშირშია იურიდიულ პირთა ცნობარში, მისი რეგისტრაციის უნიკალურ ნომერთან;

2. მრავალი-ერთთან (M :1) – განხილული ატრიბუტი შეიძლება წარმოადგენდეს რომელიმე ცნობარის ელემენტს, რომელიც შესაბამისად არის მნიშვნელოვანი ატრიბუტი. მაგ., ორგანიზაციულ – ტერიტორიული კავშირის აღწერა, რომელშიც წარმოდგენილია ორგანიზაცია, თავისი უნიკალური კოდით, ხოლო ტერიტორია შეიძლება წარმოვადგინოთ, როგორც თეორიულად შეუზღუდავი ორგანიზაციათა სიმრავლე. ე.ი. ატრიბუტი „ტერიტორია“ აღწერს ორგანიზაციებსაც;

3. ერთი-მრავალთან (1:M) – ატრიბუტი წარმოადგენს ეგზემპლიარების განუსაზღვრელ რაოდენობას, ხოლო განხილული ატრიბუტის ეგზემპლიარი დაკავშირებულია მხოლოდ ერთ იერარქიულად მაღლა მდგომ ეგზემპლიართან. მაგ., საბაზო ანგარიშებსა და ერთ მესაკუთრეს (იურიდიულ პირს) შორის კავშირი;

4. მრავალი-მრავალთან (M:N) – განსახილველი ატრიბუტები ერთმანეთს უკავშირდება მრავალი საშუალებით, რაც რომელიმე ერთ ეგზემპლარს საშუალებას აძლევს დაუკავშირდეს ნებისმიერ დანარჩენს. მაგ., იურიდიული პირების და დამფუძნებელთა სია. ყოველი იურიდიული პირი შეიძლება იყოს დამფუძნებელი და ყოველი დამფუძნებელი - იურიდიული პირი.

მონაცემთა საცავი ორიენტირებულია განსაზღვრულ საგნობრივ სფეროზე და ორგანიზებულია მონაცემთა ოპერატიული ბაზიდან შემოსულ სტრუქტურულად გადამუშავებულ მონაცემთა ქვესიმრავლების საფუძველზე. ინფორმაციის წყაროს წარმოადგენს სხვადასხვა ორგანიზაციათა დანართები (აპლიკაციები), რომლებიც გამოიყენებს განსხვავებულ პროგრამულ პლატფორმებს და უკავშირდება ოპერატიულ მონაცემთა ბაზას ინტერნეტის საშუალებით (on-line რეჟიმი). შესაძლებელია აგრეთვე სხვა სახის კავშირების (off-line რეჟიმი) გამოყენებაც.

მონაცემთა საცავში ინახება მონაცემთა სტრიქონების არა მთელი სიმრავლე, არამედ ამა თუ იმ ხარისხით გაერთიანებული (აგრეგირებული) ინფორმაცია, რაც ხელს უწყობს მესსიერების ეფექტურად გამოყენებას.

დღეს აზრთა სხვადასხვაობას იწვევს ის საკითხი, თუ რა განსხვავებაა მონაცემთა საცავს, მონაცემთა ბაზასა და ოპერაციულ სისტემას შორის. ამ განსხვავების ასახსნელად მნიშვნელოვანია თვით მონაცემთა საცავის ფუნქციური ხასიათის გაგება.

ნებისმიერი ოპერაციული სისტემა ან მონაცემთა ბაზა ორიენტირებულია კომპიუტერული რესურსების მართვის ან მონაცემთა დამუშავების ოპერაციებზე, მაშინ როდესაც მონაცემთა საცავში ყველაზე მნიშვნელოვანია თვით საინფორმაციო ობიექტი და მისი მიზნობრივად დამუშავებული შედეგების მომხმარებელზე მიწოდების მოხერხებულობა. საცავის ოპერატიულ მონაცემთა ბაზაში ინახება მხოლოდ აქტუალური (ახლად შემოსული) მონაცემები, მაშინ როდესაც საცავში და არქივში თავმოყრილია, კომპანიების მუშაობის ისტორიის მთელი საინფორმაციო სპექტრი. ისტორიული მონაცემების შენახვა არის ერთ-ერთი აუცილებელი პირობა და იგი წარმოადგენს მონაცემთა საცავის მთავარ ღირსებას, რადგან ამ

ინფორმაციის დახმარებითაა შესაძლებელი სრულყოფილ ანალიზის წარმოება.

საცავში მონაცემები ინახება ცალკეული ფაქტების ცხრილების სახით, რომელთა განხილვაც შესაძლებელია სხვადასხვა კუთხით.

საბაზო სისტემების მონაცემთა საცავში ინფორმაცია თავმოყრილია არა მხოლოდ ანალიზისათვის. შეიძლება ინფორმაცია მივიღოთ საბაზო საქმიანობის ისეთ ფუნქციურ საკითხებზე, როგორცაა საბუღალტრო საქმე, ბანკის მართვა, სახაზინო საქმე, კადრების მართვა და ა.შ.

საცავში არსებული ინფორმაციის საფუძველზე გადასაწყვეტი საკითხები მონაცემთა ანალიზის ინსტრუმენტის დახმარებით შეიძლება განხილულ იქნეს სხვადასხვა დონეზე, დაწყებული გაერთიანებული ანგარიშების შემადგენლობიდან, დამთავრებული ბიუჯეტით და ბანკის საქმიანობის ანალიზით.

კლიენტთა ბაზის და საბაზო პროდუქტების შეფასების საშუალება, მათი სტრუქტურული და დინამიკური ცვლილების ანალიზი, აგრეთვე აქტივის და პასივის ხარისხის შეფასება და ბანკებში მიმდინარე ბიზნეს-ოპერაციები, საშუალებას გვაძლევს დროულად შევცვალოთ ტაქტიკური და სტრატეგიული გადაწყვეტილებები.

მაგალითად, მონაცემთა საცავის ანალიზის დახმარებით შესაძლებელია გადავჭრათ ისეთი რთული ამოცანა, როგორცაა სხვადასხვა ობიექტების ეკონომიკური უსაფრთხოების (იგულისხმება გაკოტრება) დაცვა.

2.2. მონაცემთა საცავის დაპროექტების ამთოლურ-ორბანიზაციული საფუძვლები

მონაცემთა საცავის ცნება აქტუალური გახდა 2000 წლიდან და იგი დღემდე, საინფორმაციო სიტემების დაპროექტებაში, პერსპექტიულ მიმართულებად ითვლება. საწყის ეტაპზე საინფორმაციო სისტემების დაპროექტება ვითარდებოდა „ინფოლოგიკური“ გზით, ხოლო ახალი მიმართულების განვითარების ტენდენციამ ფართო გამოვლინება ჰპოვა „მონაცემთა საცავის“ სახით. ინფორმატიკაში როგორც ბევრი ტერმინი, ასევე ეს ცნებაც

მოკლებულია ზუსტ განმარტებას, რაც პრაქტიკაში წარმოქმნის შეუთავსებლობას. არსებობს განმარტების მრავალგვარი ახსნა რომელთაგან ერთ-ერთი ასეთია:

მონაცემთა საცავი არის რომელიმე ორგანიზაციისთვის განკუთვნილი სპეციალური ბაზა, სადაც მიმდინარე ოპერატიული სამუშაოს შესრულებისას თავს იყრის ქრონოლოგიურ ინფორმაციათა მთელი სპექტრი, რომელთა დანიშნულებაცაა მომხმარებელს მიაწოდოს ინტერნეტ გვერდებზე მიზნობრივად განლაგებული საინფორმაციო ბლოკები.

ასეთი სახის მონაცემთა ბაზის გამოყენება იმ შემთხვევაშია შესაძლებელი, თუ უზრუნველყოფილ იქნება შემდეგ ფუნქციათა შესრულება:

F1() - მონაცემთა შერჩევა.

„მონაცემთა მოხვედრა საცავში“- არის მთელი რიგი პროცესებისა, რომელიც ემყარება გარკვეულ კანონზომიერებას. პირველ რიგში ერთ-ერთი მთავარი პირობაა, რომ მონაცემები საცავში უნდა მოხვდეს საჭირო სახით, ხოლო მეორე რიგში უნდა ხდებოდეს ამ მონაცემთა რეგულირება საჭიროებისამებრ. „საჭირო სახეში“ იგულისხმება, როგორც მინიმუმ, ინფორმაცია უნდა მოვიყვანოთ საჭირო ფორმატში. ერთი და იმავე სახის მონაცემი შეიძლება მივიღოთ სხვადასხვა წყაროებიდან განსხვავებული ფორმატით, რაც წარმოქმნის შეუთავსებლობას, ამიტომ უნდა მოხდეს დასახელებათა უნიფიცირება.

მონაცემთა დამუშავების კიდევ ერთ სახეს წარმოადგენს მონაცემთა პირველადი გადამუშავება. საცავში წარმოდგენილ მონაცემებს თუ სიღრმისეულად დაუკვირდებით, შევნიშნავთ რომ სხვადასხვა კონტექსტში განხილვისას ისინი იძენს სხვადასხვა მნიშვნელობებს, რამაც შეიძლება მიგვიყვანოს არასასურველ შედეგებამდე. პირველადი დამუშავებისას ხდება შეცდომების გამომჟღავნება სტატისტიკური დამუშავების სხვადასხვა საშუალებებით, შესაძლებელია უკვე გამოყენებულ მონაცემთა ჩამოცილება და აგრეთვე გამოტოვებულ მონაცემთა აღდგენა. ოპერატიული დამუშავების სისტემიდან წამოსული მონაცემები არ

არის ერთჯერადი გამოყენების, ამიტომ მისი ნორმალური ფუნქციონირებისათვის უნდა მოქმედებდეს საცავში მონაცემთა გადაცემის პროცედურის შემსრულებელი პროგრამები და მათი პირველადი დამუშავება კავშირში უნდა იყოს გარე მოვლენებთან.

F2() - მონაცემთა ურთიერთშეთანხმება.

როგორც ყველა მონაცემთა ბაზა, მონაცემთა საცავიც შეიძლება განაწილებული იყოს კომპიუტერული ქსელის კვანძებში, სადაც მოთავსებული ინფორმაცია მიიღება განსხვავებული წყაროებიდან. იმისათვის, რომ ურთიერთშეთანხმებულად მოქმედებდეს ინფორმაციის მოძვლადეული სხვადასხვა წყაროები, მონაცემთა საცავის ფუნქციონირების ქვესისტემა აუცილებელია სარგებლობდეს მონაცემთა სტრუქტურის აღწერით. ასეთი აღწერისათვის იყენებენ ლექსიკონ-ცნობარს, სადაც თავმოყრილია ცნობები – მონაცემთა ფორმატზე, სტრუქტურაზე, ინფორმაციის მიმღებ არსებზე, წყაროებზე და ა.შ.

F3() - მონაცემთა მოპოვება.

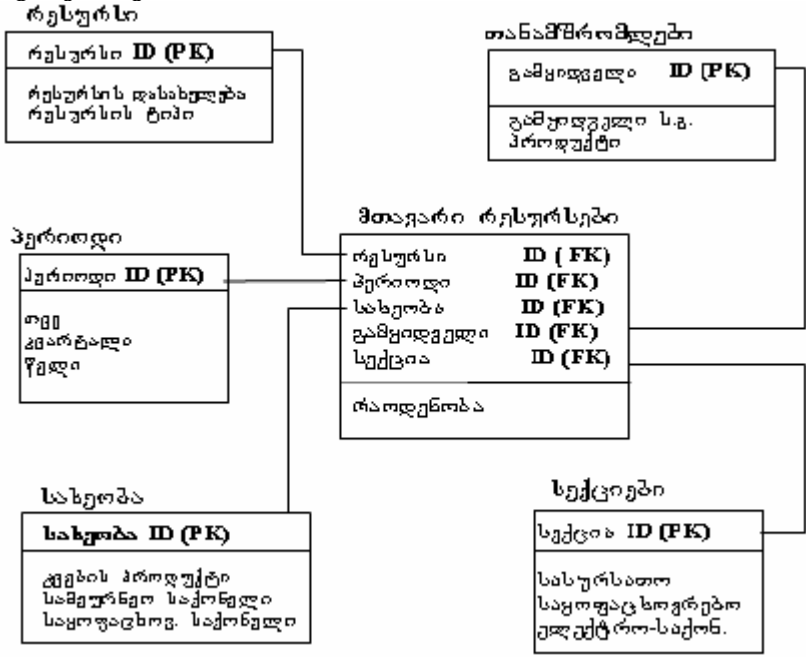
მონაცემთა საცავსა და ოპერატიულ მონაცემთა ბაზას შორის არსებული განსხვავების შესაბამისად, არსებობს მათთან მიდგომის განსხვავებული საშუალებებიც. მონაცემთა საცავთან დაკავშირებისთვის გამოიყენება არარეგლამენტირებულ მოთხოვნათა სისტემა (ad hoc query). ეს სისტემა მოთხოვნათა ფორმულირებისას გამოირჩევა მოქნილობით იგი საშუალებას გვაძლევს მონაცემებზე ანალიზის ჩატარებისას ვისარგებლოთ როგორც ზოგადი, ასევე მათი დეტალური ანალიზით.

ანალიზის მეორე სახე, რომელიც გამოიყენება ასეთ სისტემებში, გათვლის შესაძლო შემთხვევებს (what-if-analysis), როგორცაა: რეალურ მონაცემებში, რამდენიმე მაჩვენებლის შეცვლა და მათი ანალიზი, რის შედეგადაც იცვლება საერთო მდგომარეობა. ზოგადად, არარეგლამენტირებულ მოთხოვნათა სისტემის მომსახურებისთვის შეიძლება გამოვიყენოთ “მონაცემთა დამუშავების ოპერატიულ ანალიზური სისტემა“ (OLAP) [30,32,38].

მონაცემთა საცავი ეყრდნობა რელაციურ და მრავალგანზომილებიან ბაზებს. პირველადი სახის მონაცემები ინახება

რელაციურ მონაცემთა ბაზაში (უფრო ზუსტად რელაციურ-ორიენტირებულში) და განთავსებულია სპეციფიურ სქემაში, რომელსაც „რადიალური“ სქემა ან (star schema), ეწოდება.

ვ. ინმონი მონაცემთა საცავის კონკრეტულ სტრუქტურას თითქმის არ განიხილავს, ხოლო რელაციური ბაზების ფუძემდებელმა, ე.ფ. კოდმა წარმოადგინა ოპტიმიზირებული მონაცემთა ბაზების სქემა [31,33,37]. მას გააჩნია ერთი მთავარი ცხრილი, რომელთანაც დაკავშირებულია რამდენიმე დამოკიდებული ცხრილი, მთავარი ცხრილის საშუალებით ხდება დამოკიდებულ ცხრილთა მონაცემებით მომარაგება. მაგ., კომერციული ობიექტის სტრუქტურის რადიალური სქემა (ნახ.2.2). მთავარი ცხრილი ანუ ფაქტების ცხრილი ასახავს კვლევის ობიექტს, იგი შესაძლოა შედგებოდეს მილიონობით სტრიქონისგან.



ნახ.2.2

მონაცემთა ორგანიზების განსხვავებული მიდგომაა მონაცემთა საცავში, კერძოდ მრავალგანზომილებიანი სისტემის გამოყენება. აქ მონაცემები ინახება არა ცხრილში არამედ მრავალგანზომილებიან

კუბში. ჰიპერკუბის მოდელი მომხმარებლისთვის არის თვალსაჩინო და OLAP ანალიზისთვის ხელსაყრელი.

მონაცემთა მრავალგანზომილებიან წარმოდგენას ახლავს გარკვეული სირთულეებიც, ინფორმაციის დამუშავების მეთოდოლოგიის შერჩევის თვალსაზრისით [39].

F4() - კლასიფიკაცია.

ეს არის მონაცემთა კვლევის ერთ-ერთი ცნობილი მეთოდი, სადაც განიხილება მონაცემთა ტიპები (კლასები). კლასი არის ერთი ტიპის ობიექტების სიმრავლე, რომელთაც აქვთ მსგავსი სტრუქტურა და ქცევა [33]. მათი საშუალებით შეგვიძლია განვსაზღვროთ მოვლენები, სიტუაციები და პროცესები.

ობიექტები (ობიექტი განიხილება, როგორც გარკვეული არსი რომელიც ხასიათდება მდგომარეობით და ქცევით) უნდა აღიწეროს შემდეგი სიდიდეებით, როგორიცაა: სიმპტომები, მაჩვენებლები, პარამეტრები. ინფორმაცია ყოველი კლასის შესახებ მოცემულია ობიექტების დახმარებით (დაკვირვება, პრეცედენტი) აქედან გამომდინარე შეგვიძლია განვსაზღვროთ, ობიექტთა გარკვეულ ჯგუფში თუ რომელ კლასს მიეკუთვნება.

უნდა ვიპოვოთ კრიტერიუმები, რათა განვსაზღვროთ თუ რომელი კლასიფიკაციის კატეგორიას მივაკუთვნოთ ობიექტი. კრიტერიუმის შერჩევა ეყრდნობა უკვე კლასიფიცირებული ობიექტის ხასიათის შესწავლას.

ამოცნობის ალგორითმი გამოიყენება მედიცინაშიც, როდესაც გარკვეული სიმპტომებით და ამბულატორიული გამოკვლევის მონაცემებით ხდება დიაგნოზის დიაგნოსტიკა, აგრეთვე ტექნიკაში, როდესაც მაკონტროლებელი მოწყობილობის მაჩვენებლებით და ექსპერტული მონაცემების შეჯერებით შეგვიძლია დავადგინოთ ოპტიმალური კრიტერიუმები [37].

ბიზნესის ბევრ სფეროში შეინიშნება მუდმივი დამკვეთის დაკარგვის პრობლემა. კლასიფიკაციის ინსტრუმენტის გამოყენებით შესაძლოა გამოვყოთ დამკვეთის მოთხოვნათა ხასიათის შესაბამისი მონაცემთა ჯგუფი [32]. სავარაუდოა, რომ წინასწარ შექმნილ მოდელს გამოუჩნდება სასურველი დამკვეთი, ასეთი სახის მოდელი შესაძლოა იყოს დამკვეთზე ზემოქმედების ისეთი საშუალება,

როგორცაა რეკლამა, იგი საშუალებას იძლევა სხვადასხვა კატეგორიის დამკვეთის მოსაზიდად.

F5() – კლასტერიზაცია.

კლასტერიზაცია მოგვაგონებს კლასიფიკაციის ფორმას, იმ განსხვავებით, რომ კლასიფიკაციის კრიტერიუმები არ არის მოცემული. კლასტერიზაცია მონაცემთა კვლევისას საშუალებას გვაძლევს აღმოვაჩინოთ მონაცემები, რომლებიც დაჯგუფებულია რაიმე ნიშან თვისებების მიხედვით ისე, რომ ერთ ჯგუფში გაერთიანებული მონაცემები „მსგავსია“, ხოლო სხვადასხვა ჯგუფში „არაა მსგავსი“.

კლასტერიზაციის ალგორითმი, როგორც პირველადი ანალიზის ინსტრუმენტი, შეუცვლელია მრავალგანზომილებიან მონაცემთა ჯგუფის დამუშავებისას, ამასთან იგი ძალიან ეკონომიურია [38].

F6() – ოპერატიული ანალიზი OLAP-ინსტრუმენტი.

მონაცემთა საცავების დაპროექტებისა და მისი ფუნქციონირებისათვის, მეთოდური თვალსაზრისით ჩვენ ვიყენებთ ე.ფ. კოდის მიერ ჩამოყალიბებულ პრინციპებს[61]. ესაა ის 12 წესი, რომელსაც უნდა აკმაყოფილებდეს ნებისმიერი განაწილებული სისტემა მონაცემთა საცავით, რათა ჩატარდეს საინფორმაციო ბლოკების სრულფასოვანი ოპერატიული ანალიზური სამუშაოები.

1. მონაცემთა მრავალგანზომილებიანი კონცეპტუალური წარმოდგენა. მომხმარებელ-ანალიტიკოსი საპრობლემო სფეროს, თავისი ბუნების მიხედვით ხედავს როგორც მრავალგანზომილებიანს. შესაბამისად OLAP- მოდელიც უნდა იყოს მრავალგანზომილებიანი. ასეთი ტიპის კონცეპტუალური სქემა (მომხმარებელთა წარმოდგენები) აიოლებს მოდელირებას, ანალიზს და გამოთვლებს;

2. გამჭვირვალობა. OLAP წარმოდგენილი უნდა იყოს ღია არქიტექტურის კონტექსტში, სადაც მომხმარებელს საშუალება ექნება დროის ნებისმიერ მონაკვეთში, ანალიზური ინსტრუმენტის საშუალებით დაუკავშირდეს სერვერს და მიიღოს მისთვის სასურველი ინფორმაცია;

3. მიღწევადობა. OLAP – ის მომხმარებელ ანალიზატორს უნდა ჰქონდეს ანალიზის ჩატარების საშუალება, რომელიც ემყარება საერთო კონცეპტუალურ სქემას, რომელშიც განთავსებულია რელაციური მონაცემთა ბაზა სწარმოთა შესახებ არსებული ყველა

ახალი და ძველი მონაცემებით. ეს ნიშნავს, რომ OLAP – მა უნდა წარმოადგინოს თვისი საკუთარი ლოგიკური სქემა, რათა შეასრულოს შესაბამისი გარდაქმნა და მომხმარებელს წარუდგინოს მონაცემები. გარდა ამისა აუცილებელია წინასწარ იმაზე ზრუნვა, თუ სად, როგორ და როგორი სახის ფიზიკური ორგანიზაციის მონაცემები იქნას გამოყენებული. OLAP სისტემამ უნდა შეასრულოს ისეთი მონაცემების დამუშავება, რომელთა მოთხოვნაც რეალურად არსებობს;

4. ანგარიშთა დამუშავებისას მუდმივი წარმადობა. თუ ანალიტიკოსის მიერ ჩატარებული გაზომვათა რაოდენობა ან მონაცემთა ბაზების რიცხვი მნიშვნელოვნად იზრდება, მომხმარებელ ანალიზატორისთვის ეს პროცესი უნდა დარჩეს შეუმჩნეველი და არ უნდა აისახებოდეს საწარმოო პროცესების წარმადობის შემცირებაზე;

5. კლიენტ სერვერის – არქიტექტურა. მონაცემთა დიდი ნაწილი, რომელიც მოითხოვს ოპერატიულ ანალიზურ გადაამუშავებას, უნდა მუშაობდეს კლიენტ-სერვერულ რეჟიმში. ამ თვალსაზრისით აუცილებელია, რომ ანალიზური ინსტრუმენტის სერვერული კომპონენტები იყოს „ინტელექტუალური“, რადგან განსხვავებულ კლიენტებს შეეძლოთ დაუკავშირდნენ სერვერს და გამოიყენონ პროგრამიული პაკეტები. „ინტელექტუალურ“ სერვერს უნდა შეეძლოს მონაცემთა ბაზის შეუთავსებადი ლოგიკური და ფიზიკური სქემის ასახვა და გაერთიანება. ეს უზრუნველყოფს გამჭვირვალებას და საშუალებას იძლევა აიგოს საერთო კონცეპტუალური, ლოგიკური და ფიზიკური სქემები;

6. მრავალგანზომილება. გაზომვის ყოველი მცდელობის დროს გამოყენებული უნდა იყოს მიუკერძოებელი სტრუქტურა და ოპერაციული შესაძლებლობა. დამატებითი ოპერაციული შესაძლებლობა უნდა მიეცეს ერთ-ერთ რომელიმე ცდას და თუ ეს გაზომვა სიმეტრიული იქნება სხვა გაზომვის შედეგების, მაშინ ცალკე აღებული ფუნქცია შეიძლება წარმოვადგინოთ ნებისმიერი გაზომვის სახით;

7. დინამიკური მართვა გამონთავისუფლებული რეჟიმით. OLAP – ინსტრუმენტის ფიზიკური სქემა უნდა ადაპტირდებოდეს სპეციფიკურ ანალიტიკურ მოდელთან, რათა ოპტიმალურად მართოს გამონთავისუფლებული მატრიცა. ერთი დაცლილი მატრიცისთვის არსებობს ერთადერთი ოპტიმალური ფიზიკური სქემა. **OLAP** –

ინსტრუმენტის ბაზური ფიზიკური მონაცემები პრაქტიკული ოპერაციებისთვის, რომელთაც აქვთ დიდი ანალიზური მოდელი უნდა კონფიგურირდებოდეს ნებისმიერ ქვესიმრავლესთან. თუ **OLAP** – ინსტრუმენტს არ შეუძლია გააკონტროლოს და დაარეგულიროს საანალიზებელი მონაცემების მნიშვნელობები, ის ჩაითვლება უსარგებლოდ და არასაიმედოდ;

8. მრავალმომხმარებლობა. ხშირ შემთხვევაში მომხმარებელ-ანალიტიკოსი დასმულ მოთხოვნებს აყენებს ერთ ანალიზურ მოდელთან ან ქმნის განსხვავებულ მოდელს ერთი სახის მონაცემებიდან. **OLAP** – ინსტრუმენტი კი საშუალებას გვაძლევს მივიღოთ უსაფრთხო, სრულყოფილი და ზუსტი ანალიზური შედეგები;

9. შეუზღუდავი გადამკვეთი ოპერაციები. მონაცემთა შემოწმების სხვადასხვა დონე და გაერთიანების გზა, მათი იერარქიული ბუნების გათვალისწინებით მჭიდრო კავშირშია **OLAP** - მოდელთან ან დანართთან. თვითონ ინსტრუმენტი უნდა მოიაზრებოდეს შესაბამის გამოთვლებთან და არ უნდა მოსთხოვოს მომხმარებელს თავიდან განსაზღვროს გამოთვლები და ოპერაციები. გამოთვლები მოითხოვს რომელიმე გამოყენებულ ენაში განსხვავებული ფორმულების განსაზღვრას. ასეთი ენა შეიძლება გამოვიყენოთ ნებისმიერი სიდიდის მონაცემთა მანიპულირებისთვის და არ შეზღუდოს მონაცემები არსებული კუბის უჯრედებს შორის და კონკრეტული უჯრედების საერთო ატრიბუტებზე;

10. მონაცემთა ინტუიციური მანიპულაცია. მონაცემთა დეტალიზაციის, გაერთიანების და სხვა მანიპულაციები უნდა იყენებდეს ცალკეულ უჯრედებზე ანალიზური მოდელის შედეგებს და არ უნდა იყენებდეს მომხმარებლის ინტერფეისებს. მომხმარებელ ანალიტიკოსს უნდა ჰქონდეს ყველა აუცილებელი პირობა იმისა, რომ მიიღოს სრულყოფილი ინფორმაცია;

11. ანგარიშების მიღების მოქნილი საშუალება. შეტყობინებათა დამუშავება და პასუხის გაცემა უნდა იყოს მოქნილი და ელასტიური. მომხმარებელს უნდა შეეძლოს მონაცემთა კომბინირება და გაანალიზება. მოქნილობა მნიშვნელოვანია, რათა ყურადღება გამახვილდეს მონაცემთა განმასხვავებელ ნიშნებზე. თუ რაიმე სირთულე წარმოიქმნება, უნდა შევჩერდეთ ინდივიდუალურ

ინფორმაციულ მოთხოვნაზე და შეირჩეს მხოლოდ მომლოდინე მოთხოვნა;

12. შეუზღუდავი ზომები და აგრეგაციათა რაოდენობა.

გამოკვლევებმა აჩვენეს, რომ აუცილებელი გაზომვა ერთდროულად შეიძლება ჩატარდეს 19-ჯერ. აქედან გამომდინარე შეიძლება ვთქვათ, რომ ანალიზური ინსტრუმენტი საშუალებას გვაძლევს ერთდროულად ვაწარმოოთ 15-დან 20-მდე გაზომვა, ამასთან თითოეული გაზომვის მცდელობა არ არის შემოსაზღვრული დადგენილი რიცხვით.

ეს პირობები შეიძლება ჩავთვალოთ, ოპერატიული ანალიზური დამუშავების თეორიულ ბაზისად. როგორც უკვე ავლინებთ **OLAP**-ში ძვეს მონაცემთა დამუშავების მრავალგანზომილებიანი სტრუქტურის იდეა. როდესაც ვსაუბრობთ **OLAP**-ზე, უნდა ვიგულისხმობთ, რომ ეს არის მონაცემთა ლოგიკური სტრუქტურის მრავალგანზომილებიანი ანალიზური ინსტრუმენტი.

2.3. მონაცემთა ტიპიზაციის, კლასიფიკაციისა და კატალოგიზაციის ამოცანის გადაწყვეტა

საავტომატიზაციო ობიექტის განაწილებული სისტემის მონაცემთა საცავში განსათავსებელი საინფორმაციო ბლოკები შეივსება ოპერატიულ მონაცემთა ბაზებიდან, მათი წინასწარი დამუშავების საფუძველზე (მეთოდური საფუძვლები წინა პარაგრაფში იყო განხილული). როგორც ვიცით, ოპერატიულ ბაზაში საწყისი მონაცემები თავს იყრის გარე ორგანიზაციებიდან ინტერნეტის საშუალებით, ან სხვა სახის კომუნიკაციებიდან.

საინფორმაციო ბლოკები, ან უფრო ზოგადად, მონაცემთა მანქანური ფაილები განსხვავებული ტიპებისაა: ტექსტები და სუპერტექსტები (H), ცხრილები (T), გრაფიკები (G), აუდიო (A) და ვიზუალური (V) მასალა და ა.შ.

მათემატიკური მოდელი შეიძლება ჩავწეროთ ტეტრადით:

$$B = \langle R, M, S, F \rangle, \text{ სადაც} \quad (2.1)$$

R – საცავის კლასიფიცირებული ობიექტებია.

$R = \{ \}$, . აქ i კლასიფიკაციის ტიპია და იგი შეესაბამება ზემოაღნიშნულ ცვლადებს: {H, T, G, A, V}.

M – მონაცემთა ბლოკების მეტაინფორმაციაა, განთავსებული სამდონიან იერარქიულ კატალოგსა (C_R , C_A , C_V) და ინდექსურ ცხრილებში (T_I).

$$M = \{C_R, C_A, C_V, T_I\}. \text{ სადაც} \quad (2.2)$$

C_R - საინფორმაციო ბლოკების აღწერის რელაციათა სიმრავლეა: $C_R = \{ \}$, სადაც j = რელაციათა ინდექსებია (რელაციათა სიმრავლის სიმძლავრე).

C_A - საინფორმაციო ბლოკების აღწერის ატრიბუტების სიმრავლეა: $C_A = \{ \}$, სადაც k = ატრიბუტების ინდექსებია (ატრიბუტთა სიმრავლის სიმძლავრე).

C_V - ატრიბუტების მნიშვნელობათა სიმრავლეა:

$C_V = \{ \}$, სადაც $k = A_k$ -ატრიბუტის ინდექსები, ხოლო z ამ ატრიბუტის შესაბამისი მნიშვნელობათა ინდექსებია (ატრიბუტის ქვესიმრავლის სიმძლავრე).

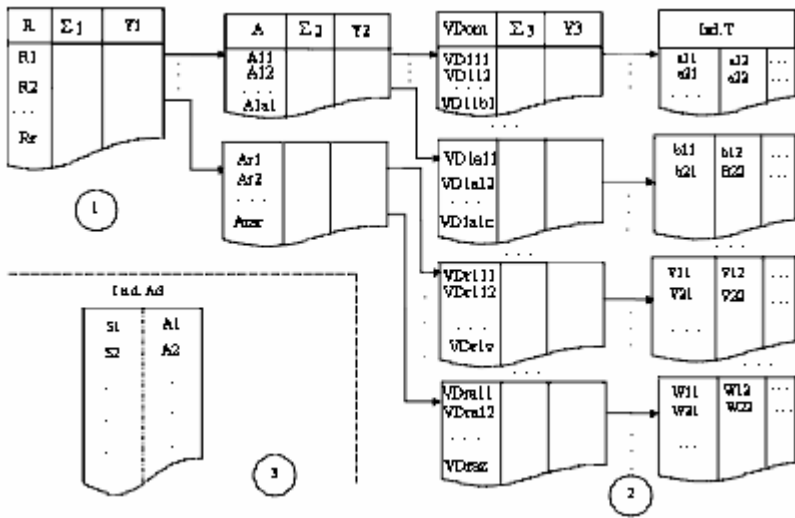
T_I - ინდექსების ცხრილის კორტეჟების (ან სტრიქონების, ჩანაწერების) სიმრავლეა: $T_I = \{ \}$, სადაც j = რელაციათა ინდექსები, $k = A_k$ -ატრიბუტის ინდექსები, ხოლო V_z ამ ატრიბუტის შესაბამისი მნიშვნელობათა ინდექსებია. ინდექსების ცხრილის სვეტების რაოდენობაა n , ხოლო სტრიქონების რაოდენობა - m .

S – საცავის მეხსიერების ცენტრალური ცხრილია, რომელთა კორტეჟებშიც მოთავსებულია საინფორმაციო ბლოკების ფიზიკური მისამართები და ინდექსური ცხრილების ელემენტთა მნიშვნელობები. ამგვარად, მყარდება ლოგიკური კავშირი მოთხოვნაში დასახელებულ j -ურ რელაციის k -ურ ატრიბუტის z -ურ მნიშვნელობასა და მეხსიერების ფიზიკურ a -ურ მისამართს შორის.

F – მონაცემთა საცავის ელემენტების (R), ანუ კლასიფიცირებული ობიექტების დამუშავების ოპერაციებია (კლასთა მეთოდებია). ჩვენ წინა პარაგრაფში განვიხილეთ ზოგიერთი ძირითადი ოპერაცია, მაგალითად: $F1()$ - მონაცემთა შერჩევა, $F2()$ - მონაცემთა ურთიერთშეთანხმება, $F3()$ - მონაცემთა მოპოვება, $F4()$ –

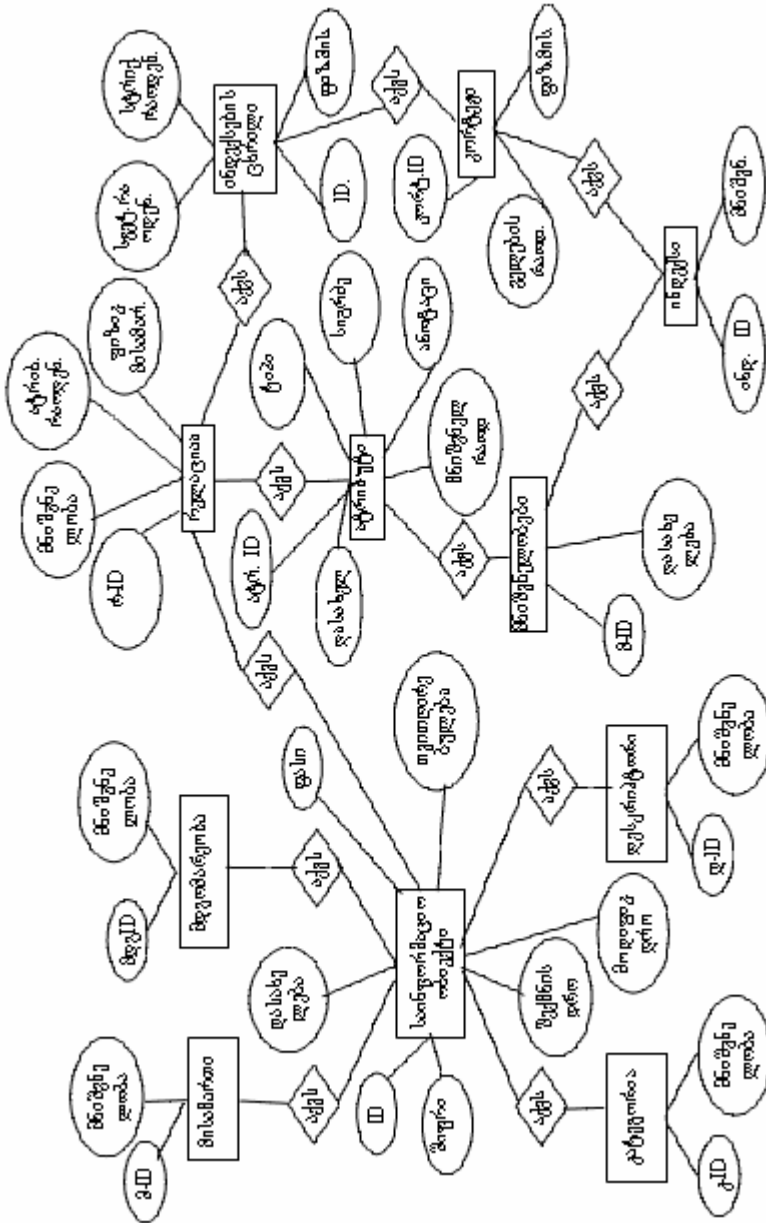
კლასიფიკაცია, F5() –კლასტერიზაცია, F6() – ოპერატიული ანალიზი OLAP-ინსტრუმენტით და სხვ. აქ ემატება მონაცემთა ტრანსფორმაციის, კონვერტირების და რესტრუქტურის ფუნქციებიც. მომხმარებელთა მოთხოვნების დამუშავების ამოცანა განიხილება 2.4 პარაგრაფში.

2.3 ნახაზზე მოცემულია მონაცემთა საცავის მეტაინფორმაციის ორგანიზებისა და მართვის სამლონიანი კატალოგისა და ინდექსური ცხრილების გრაფიკული წარმოდგენა, 2.4-ზე კი ილუსტრირებულია საინფორმაციო ობიექტების დეტალური აღწერის კონცეპტუალური მოდელი არსთა-დამოკიდებულების (ER) დიაგრამის საშუალებით.



ნახ.2.3

პროგრამული რეალიზაციის საკითხი დაკავშირებულია მონაცემთა განაწილებული ბაზების მართვის სისტემის, მაგალითად, Ms SQL Server-2000-ის გამოყენებასთან.



სსს.2.4

პირველ ეტაპზე საჭიროა დაპროექტდეს მონაცემთა საცავის საინფორმაციო ობიექტების მეტაინფორმაციის სტრუქტურა, რომელიც საერთო იქნება ყველა ტიპის (H, T, G, A, V) კლასისათვის. ქვემოთ მოცემულია ამ სტრუქტურის ცხრილი შესაბამისი ატრიბუტებითა და განმარტებებით:

ინფორმაციულ ობიექტთა მეტაინფორმაციის სტრუქტურა. ცხრ.2.1

№	ატრიბუტი	დასახელება	ტიპი	შენიშვნა
1.	ObjID	ობიექტის იდენტიფიკატორი	int	უნიკალური
2.	Shifri	ობიექტის შიფრი	varchar	
3.	Name	ობიექტის დასახელება	varchar	
4.	Category	ობიექტის კატეგორია	smallint	მეორადი გასაღები
5.	StateCode	მდგომარეობის კოდი	smallint	მეორადი გასაღები
6.	Summary	ობიექტის ანოტაცია	varchar	MEMO-ველი
7.	DescrID	ძირითად დესკრიპტორთა იდ.	smallint	მეორადი გასაღები
8.	FileSize	ფაილის ზომა ბაიტებში	int	
9.	CreatDate	ობიექტის შექმნის თარიღი	Date	
10.	ModifDate	ბოლო მოდიფიკაციის თარიღი	Date	
11.	PriceA	თვითღირებულება	Float	შენახვის ხარჯი
12.	PriceB	ფასი	Float	გასაყიდი ფასი
13.	PhysAddrID	ფიზიკური მისამართის იდენტიფ.	char	მეორადი გასაღები

2.4. მოთხოვნების წინმსწრები ანალიზისა და ტრანზაქციების სინქრონიზაციის სერიალიზაციის ამოცანის გადაწყვეტა

დაპროგრამების ობიექტ-ორიენტირებული მეთოდისა და უნიფიცირებული მოდელირების ენის (UML) საფუძველზე კლასი, თავისი მონაცემ-წევრებითა (ატრიბუტები) და ფუნქცია-წევრებით (მეთოდები), აგრეთვე კლასთაშორისი კავშირები (Class Assotiation) ითვლება საინფორმაციო სისტემების პროგრამული უზრუნველყოფის ძირითად კომპონენტებად [65,66].

ასევე მნიშვნელოვანია საკითხი აღნიშნული მეთოდების გააქტიურების მომენტების განსაზღვრისათვის. ეს მომენტები კი დამოკიდებულია დროზე და სისტემაში შემოსულ შეტყობინებებზე (მოთხოვნებზე).

შეტყობინებათა წინმსწრები ანალიზით შესაძლებელია დადგინდეს თუ რომელი საინფორმაციო ობიექტი აინტერესებს ამა თუ იმ მომხმარებელს (ან ჯგუფს).

წინმსწრები ანალიზის ალგორითმის მიზანია დროის ერთეულში შემოსული მოთხოვნების პაკეტის (რამდენიმე შეტყობინება) დესკრიპტორული ანალიზის ჩატარება და მათი დაკმაყოფილების მიზნით აუცილებელი და საკმარისი ინფორმაციული რესურსების მოძიებისა და პასუხების მომზადების მიმდევრობით-პარალელური ოპერაციების დაგეგმვის განხორციელება.

„პაკეტის“ წინმსწრები განხილვის აზრი მდგომარეობს იმაში, რომ გამოვლინდეს სხვადასხვა შეტყობინებაში ერთი და იმავე საინფორმაციო ობიექტებზე მოთხოვნა, შემდგომ ეტაპზე საცავში განმეორებითი ძებნის ოპერაციების გამოსარაცხად. ეს ამალგებს სისტემის მწარმოებლურობას და ამცირებს მოთხოვნების დაკმაყოფილების საერთო დროს.

2.1 მათემატიკურ მოდელში შევიტანოთ მომხმარებლის მოთხოვნის, ანუ შეტყობინების ელემენტი (Q - Query). მოდელი მიიღებს შემდეგ სახეს:

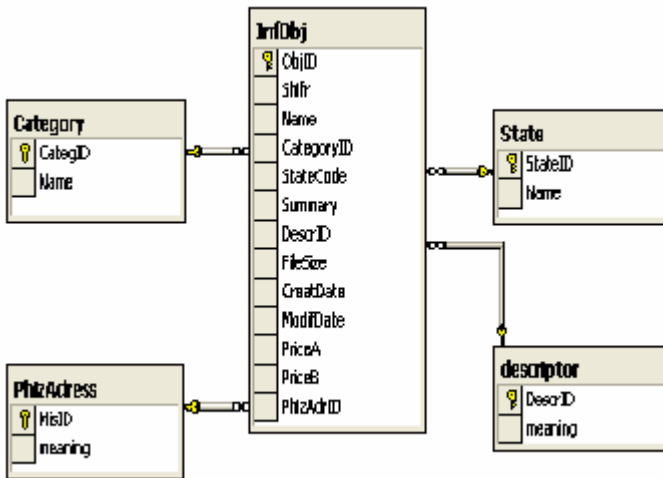
$$B = \langle R, M, S, F, Q \rangle \quad (2.3)$$

ჩვენთვის აუცილებელია ლოგიკური მექანიზმის შემუშავება, რომელიც ყოველ კონკრეტულ შეტყობინებას ცალსახად

განუსაზღვრავს მისთვის საჭირო კლასიფიცირებულ ობიექტთა სიმრავლეს მონაცემთა საცავიდან.

აღნიშნული ამოცანის გადასაწყვეტად წინასწარ საჭიროა დაძუშავდეს მონაცემთა საცავის საინფორმაციო ობიექტებისა და მათი მეტაინფორმაციის (კატალოგების) სისტემის კლასთაშორისი-კავშირების დიაგრამა.

მისი აგება შესაძლებელია წინა პარაგრაფში დაპროექტებული კონცეპტუალური მოდელის საფუძველზე. 2.5 ნახაზზე ნაჩვენებია ობიექტ-ორიენტირებული დიაგრამის კომპიუტერული რეალიზაციის ფრაგმენტი.



ნახ.2.5

ამოცანის გადაწყვეტის რეალიზაციის მიზნით ჩვენ ვიყენებთ მონაცემთა რელაციური ბაზების მართვის სისტემის **MsSQLServer_2000** პროგრამულ პაკეტს. ინფორმაციულ მონაცემთა და მათი მეტაინფორმაციის ძირითადი ცხრილები აგებულია რელაციური პრინციპით. მონაცემთა ძირითად დესკრიპტორთა სიმრავლისათვის შექმნილია ინდექსური ფაილები (სწრაფი მოძებნის ცხრილები). მომხმარებელთა შეტყობინებების ფორმირება ხორციელდება სტანდარტული SQL-ენის საფუძველზე.

მოთხოვნების დაბუღების პროცედურები (მაგ., სელექცია) ბაზებში ძირითადად შეიცავს ისეთ კრიტერიუმებს, რომელთა შესრულება ხორციელდება რელაციური ალგებრის ოპერაციებით ცხრილების კორტეჟებსა და ატრიბუტებზე.

2.3 მოდელში Q შეტყობინებათა წინმსწრები ანალიზის ჩასატარებლად ვიყენებთ:

- ე. კოდის რელაციური მოდელებისა და ალგებრის კონცეფციას:

$$R_a = \langle R, f \rangle, \text{სადაც} \quad (2.4)$$

R - რელაციური დამოკიდებულებებია მონაცემთა საცავის საინფორმაციო ობიექტების ცხრილების ატრიბუტთა სიმრავლეზე განსაზღვრული, ხოლო f - სიმრავლეთა თეორიისა და კოდის ალგებრის ოპერაციები [50].

- პეტრის ქსელების გრაფულ წარმოდგენას :

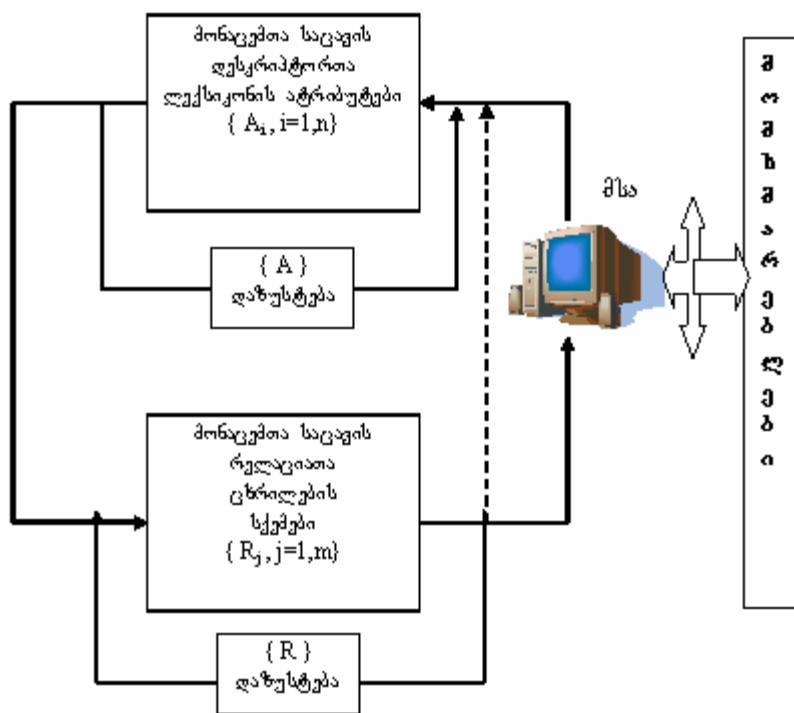
$$P = \langle S, T, I, O \rangle, \text{სადაც} \quad (2.5)$$

S - ქსელის პოზიციებია (ობიექტის ან შესაბამისი ცხრილის მდგომარეობები), T - ქსელის გადასასვლელები (რელაციური ოპერაციები გარკვეული დაყოფებით), I და O შესაბამისად არის შემავალი და გამოშვებული ფუნქციები. ქსელის პოზიციაში შეიძლება იყოს მარკერი, რომელიც ამოღებულ მავალითად, ქსელში მოთხოვნის შემოსვლას, არსებობას ან შედეგის მიღებას.

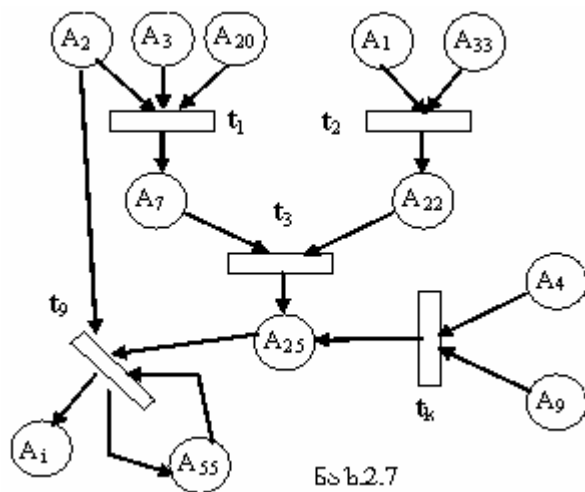
ვინაიდან მომხმარებლის მოთხოვნა პრედიკატულ ფორმაში ყალიბდება (SQL სტანდარტი), მასში ძირითადად ლოგიკური და რელაციური პროცედურები სჭარბობს.

რა თქმა უნდა მოთხოვნაში შეიძლება იყოს მათემატიკური, სტატისტიკური, აგრეგაციის, დაჯგუფებისა და სხვა სახის ფუნქციები, რომლებიც პროგრამულად მეთოდების სახითაა რეალიზებული და შენახულია პროგრამულ ბიბლიოთეკაში. 2.6 ნახაზზე ნაჩვენებია მომხმარებელთა მოთხოვნების წინმსწრები ანალიზის ორეტაპიანი მართვის პროცესის სქემა.

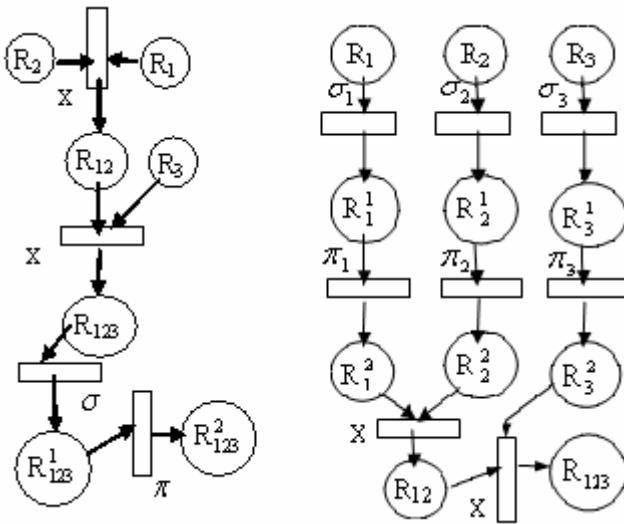
მისი შესაბამისი პეტრის ქსელის გრაფის ფრაგმენტები ატრიბუტებისა და რელაციათა სქემებისათვის წარმოდგენილია 2.7 და 2.8 ნახაზებზე.



ნახ.2.6



ნახ.2.7



ნახ.2.8

აღნიშნული საკითხის გადასაწყვეტად მულტიმომხმარებლურ რეჟიმისათვის ნაშრომში შემოთავაზებულია ტრანზაქციათა სინქრონიზაციის სერიალიზაციის ალგორითმი, რომელიც რეპოზიტორ-მენეჯერ პროგრამის კლასის ერთ-ერთ მეთოდს წარმოადგენს.

მომხმარებელთა ტრანზაქციები (საინფორმაციო ობიექტების წაკითხვა/ჩაწერის პროცედურები) იყენებს ქსელის საერთო გამოყენების რესურსებს, ამიტომაც საჭიროა მათი ბლოკირება-დებლოკირების კონფლიქტური სიტუაციების მართვა, რათა თავიდან ავიცილოთ „ჩიხური“, უსასრულო მოლოდინის მდგომარეობები[10].

ტრანზაქციათა სერიალიზაცია კი უზრუნველყოფს მათ პარალელურ შესრულებას. რომელიმე საინფორმაციო ობიექტში ცვლილებების შეტანისას ერთი ტრანზაქციით იგი ბლოკირებულია სხვა ტრანზაქციებისათვის ამ პროცესის დამთავრებამდე.

ამასთანავე, მონაცემთა საცავის მთლიანობის მოთხოვნის დაცვის მიზნით რეპოზიტორ-მენეჯერი მთლიან საცავში (კატალოგებსა და განაწილებულ ბაზებში) განახორციელებს შესაბამის ცვლილებებს.

2.5. მონაცემთა საცავის პროგრამული კომპონენტები

მონაცემთა საცავი მოიცავს ტრანსფორმაციისა და კონვერტაციის პროგრამებს, საბაზო მეტამონაცემთა სისტემას, არქივირებული შენახვის სისტემას და ინტეგრირებულ მონაცემთა საცავს [39].

მონაცემთა ბაზების სიტემა:

მონაცემთა საცავს ემსახურება ისე, როგორც სხვა მონაცემთა ბაზების სისტემები, რომელთა საშუალებითაც ხდება, მონაცემთა დამოუკიდებელ პროგრამათა ინტეგრაცია, შენახვა და მართვა. მონაცემთა ბანკების მართვის სისტემა მომხმარებელთათვის უზრუნველყოფს ადამიანსა და კომპიუტერს შორის ინფორმაციის გაცვლას. იგი შეიცავს მონაცემთა დიდ რაოდენობას, რომელთა შორისაც დამყარებულია კანონზომიერი კავშირები. მონაცემთა ბანკების მართვის სისტემა შეადგენს მონაცემთა საცავის ცენტრალურ სისტემას.

ტრანსფორმაციის პროგრამა:

ახორციელებს ინტერფეისის ფუნქციას მონაცემთა საცავსა და მონაცემთა წყაროებს შორის. მონაცემები (ინფორმაცია) განსხვავებულ მონაცემთა ბაზებიდან (იერარქიულ, რელაციურ ან ობიექტ-ორიენტირებულ) ან თანამიმდევრულ განსხვავებული ფორმატის ფაილებიდან (ASCII, ANSI, EBCDIC და ა.შ.) ექსტრადირდება. ტრანსფორმაციის წესის თანახმად ისინი ერთიანდება (Bridging), როგორც ინტეგრირებული, სუბიექტ-ორიენტირებული, მუდმივი და დროში ცვალებადი სტრუქტურები. ტრანსფორმაციის პროგრამამ უნდა უზრუნველყოს ტრანსპორტირებისათვის ფუნქციათა წარმოდგენა და აგრეთვე მონაცემთა მომზადება საცავში გადასაგზავნად. მონაცემთა საცავში შიგა მონაცემების (ინფორმაციის) უდიდესი ნაწილი შესაძლოა მიღებული იქნას განაწილებული ოპერატიული სისტემიდან. მონაცემთა საცავში ისტორიული და მიმდინარე მონაცემების შევსება, ხდება ბაზებიდან, სადაც პერიოდულად მიმდინარეობს მონაცემთა აქტუალიზაცია. თუ ოპერატიულ მონაცემთა ჩანაწერების რეგისტრაციაში მონაცემის შეტანის თარიღი უფრო ახალია, ვიდრე ბოლო ტრანსფორმაციის შეტანის დრო, მაშინ მოხდება ამ

უკანასკნელის ლიკვიდაცია. ე.ი. აუცილებლად გათვალისწინებული უნდა იქნეს ცალკეულ ტრანსფორმაციის პროცესებს შორის ვადები.

მონაცემთა გარე მიმწოდებლებია შეტყობინებათა სამსახურები, ბირჟები, პოლიტიკური საინფორმაციო სამსახურები, სამეცნიერო კვლევითი ინსტიტუტები და ბაზარი. ინფორმაციის მოწოდებისათვის გამოიყენება ისეთი საშუალებები, როგორცაა მაგალითად: Internet, CD-ROM, Flash-Memory, FD და ა. შ. ბევრი ამ მონაცემთაგან ტრანსფორმაციის პროგრამის საშუალებით, სანამ გადავა მონაცემთა საცავში უნდა წარმოდგეს გარკვეული Internet-სტანდარტული-ფორმატის სახით.

მეტამონაცემთა საბაზო სისტემა:

მეტამონაცემები არის მონაცემები - მონაცემების შესახებ. მეტამონაცემთა საბაზო სისტემა მონაცემთა საცავის შემადგენელი ნაწილია. მონაცემთა საცავში იგი ერთ-ერთ მთავარ როლს ასრულებს, მეტამონაცემების დანიშნულებაა მონაცემთა საცავში არსებული მონაცემების აღწერა და მათზე დამატებითი ცნობების მოგროვება. მეტამონაცემთა ინსტრუმენტის გამოყენება მომხმარებელს საშუალებას აძლევს აწარმოოს მონაცემთა მასივებში მანევრირება და ეხმარება მას მონაცემთა საცავში ორიენტირებისათვის.

მეტამონაცემთა საბაზო სისტემა მომხმარებელს ეხმარება მოთხოვნების შესაბამისად მონაცემთა შერჩევაში. ეს ხორციელდება მონაცემთა საცავში მეტამონაცემთა გამჭვირვალე ასახვის შედეგად.

არქივირებული შენახვის სისტემა:

უზრუნველყოფს მონაცემთა დაცვას და მათ არქივირებულ შენახვას მონაცემთა საცავში. მონაცემთა არქივირებული შენახვა, როგორც ცალკე სისტემა, მონაცემთა საცავში ამცირებს მეხსიერების უჯრედებს და ზრდის მუშაობის ეფექტურობის ხარისხს. არქივირების ეფექტური სისტემა მნიშვნელოვანია, რადგან მოკლე ვადაში შესაძლებელია მონაცემთა გადმოტვირთვა მომხმარებელთა მოთხოვნების შესაბამისად.

ხშირად სისტემაში თავს იყრის უსარგებლო ინფორმაციათა ნაკადი და იკავებს დიდ ადგილს, რაც აფერხებს სისტემის მუშაობის ეფექტურობას, არქივირებული სისტემის დანმარებით ხდება ასეთი ინფორმაციის განადგურება. შესაძლოა ასევე დეფექტური

ტრანზაქციის შედეგად მოხდეს მონაცემთა 'დაზიანება'. ამ შემთხვევაში ამოქმედდება მონაცემთა დაცვის სისტემა, რაც უზრუნველყოფს დეფექტების აღმოფხვრას და არასასურველი ინფორმაციის განადგურებას.

არქივირებული შენახვისას ყურადღება უნდა გავამახვილოთ ინფორმაციის შენახვის კანონებზე, რომელიც ითვალისწინებს არქივში ინფორმაციის შენახვის ვადებს, რადგან შენახული ინფორმაცია გარკვეული პერიოდის შემდეგ კარგავს აქტუალობას.

მონაცემთა საცავის ანალიზური დამუშავება OLAP – ტექნოლოგიის გამოყენებით.

მონაცემთა მრავალგანზომილებიან კომპლექსური ანალიზის ტექნოლოგიას უწოდებენ **OLAP (Online Analytical Processing)** ტექნოლოგიას, რომელიც ნიშნავს „მონაცემთა ოპერატიული ანალიზს“. მონაცემთა საცავში იგი წარმოადგენს მნიშვნელოვან კომპონენტს.

OLAP – ინსტრუმენტი პირველად მონაცემებს წარმოადგენს ინფორმაციის სახით, რომლის დახმარებითაც შესაძლებელი ხდება საწარმოს მოცულობის შესახებ ვიქონიოთ რეალური წარმოდგენა. ამავე დროს იგი უნიკალური ინსტრუმენტია, რომელიც საშუალებას გვაძლევს სხვადასხვა ანალიზური ჭრილით ჩავატაროთ ინფორმაციის მრავალგანზომილებიანი ანალიზი.

2.6. ოპერატიულ მონაცემთა ბაზის ინფორმაციის კონვერტირების ინსტრუმენტი

მონაცემთა საცავებში ინფორმაციული ბლოკები შეივსება ოპერატიულ მონაცემთა ბაზებიდან, სადაც გარე ორგანიზაციების საწყისი მონაცემები თავს იყრის ინტერნეტის ან სხვა სახის კომუნიკაციებიდან მოსული ინფორმაციით.

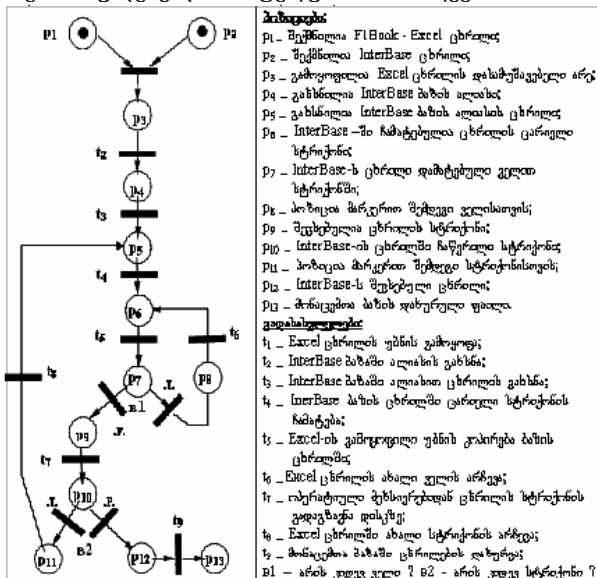
როგორც ანალიზმა გვიჩვენა, ელექტრონული ბიზნესისა და კომერციის ობიექტებზე (მაგ., დიდი სავაჭრო ცენტრები) თითქმის 60-70% გამოიყენება MsExcel-პაკეტზე აგებული ცხრილები (შეკვეთები, ფაქტურები და სხვა ანგარიშები). მონაცემთა საცავი რელაციური ტიპის ბაზებთან სამუშაოდაა ორიენტირებული, როგორებიცაა, მაგ., MsAccess, SQL Server-2000, Oracle, InterBase და ა.შ.

ამ მონაცემთა ბაზებს გააჩნიათ როგორც ერთმანეთთან, ასევე Ms Excel-თან მონაცემთა ცხრილების გაცვლის შესაძლებლობანი (Import/Export, MsQuery - ფუნქციების ან ინსტრუმენტების სახით და ODBC დრაივერით).

წინამდებარე პარაგრაფში ჩვენს მიერ შემუშავებული და წარმოდგენილია ექსცელის ცხრილების კონვერტაციის ფუნქციის ალგორითმული სქემა მონაცემთა განაწილებული ბაზების მართვის სისტემის, InterBase-ს სტრუქტურაში. ამ მონაცემთა ბაზას ფართოდ იყენებს ბორლანდის ფირმა (Borland C++ Builder და Borland Delphi საბაზო ტექნოლოგიების პროგრამული აპლიკაციებით) და ODBC დრაივერი აქ ვერ გამოიყენება. ამიტომაც ის აქტუალურად მიგვაჩნია.

ამოცანა მდგომარეობს მონაცემთა საცავის განაწილებული რელაციური ბაზის ავტომატიზებულად შესავსებად მონაცემთა პირველადი წყაროების ელექტრონული ცხრილებიდან.

2.9 ნახაზზე ნაჩვენებია აღნიშნული ამოცანის გადაწყვეტის ალგორითმის შესაბამისი პეტრის ქსელის გრაფი, რომელშიც ასახულია შესასრულებელ პროცედურათა მიმდევრობა.



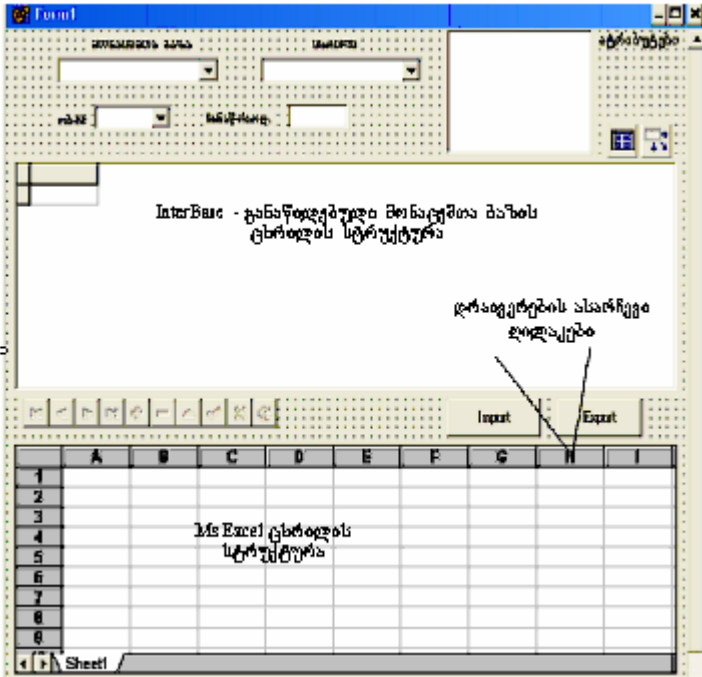
- შენიშვნა:**
- P1 - შექმნილია File-ის Excel ცხრილები
 - P2 - შექმნილია InterBase ცხრილები
 - P3 - გამოყოფილია Excel ცხრილის დასაშუქავი არე
 - P4 - გახსნილია InterBase ბაზის ალიასი
 - P5 - გახსნილია InterBase ბაზის ალიასის ცხრილები
 - P6 - InterBase-ში ჩამატებულია ცხრილის ცარიელი სტრუქტურა
 - P7 - InterBase-ს ცხრილი დაიტვირთა ველი სტრუქტურაში
 - P8 - აღივსება მარჯვნივ შექმნილი ველი
 - P9 - შევსებულია ცხრილის სტრუქტურა
 - P10 - InterBase-ის ცხრილში ჩაწერილი სტრუქტურა
 - P11 - აღივსება მარჯვნივ შექმნილი სტრუქტურისთვის
 - P12 - InterBase-ს შევსებულია ცხრილი
 - P13 - მონაცემთა ბაზის ფიზიკური ფაილი
- გადასასრულებელი**
- t1 - Excel ცხრილის უბნის გამოყოფა
 - t2 - InterBase ბაზაში ალიასის გახსნა
 - t3 - InterBase ბაზაში ალიასით ცხრილის გახსნა
 - t4 - InterBase ბაზის ცხრილში ცარიელი სტრუქტურის ჩამატება
 - t5 - Excel-ის გამოყოფილი უბნის კონვერტაცია ბაზის ცხრილში
 - t6 - Excel ცხრილის ახალი ველის არჩევა
 - t7 - ოპერატიული ბუქსიდან ცხრილის სტრუქტურის დაგეგმვა დისკზე
 - t8 - Excel ცხრილში ახალი სტრუქტურის არჩევა
 - t9 - მონაცემთა ბაზაში ცხრილის დასრულება
 - t10 - ალის კლავი ალიას ? B2 - არის კლავი სტრუქტურა ?

ნახ.29. პეტრის ქსელის გრაფი დოკუმენტის Button.Click - Import

2.10 და 2.11 ნახაზებზე ილუსტრირებულია რეალიზებული C++პროგრამის ფრაგმენტი და შესაბამისი საინტერფეისო ფანჯარა მონაცემთა ბაზის ადმინისტრატორისათვის.

სისტემა რეალიზებულია Borland C++Builder გარემოში.

```
//— Unit1.cpp — პროგრამის ტექსის ფრაგმენტი ———
int i,j, Ob_Num, Dan_Num, Rec_Num;
void __fastcall TForm1::FormCreate(TObject *Sender)
{ Session->GetAliasNames(cbAlias->Items); }
void __fastcall TForm1::cbAliasChange(TObject *Sender)
{ Session->GetTableNames(cbAlias->Text, "", true, false, cbTable->Items);
  cbTable->ItemIndex = 0;
  if (cbTable->Text == "")
  { ShowMessage("Table not selected"); return; }
  Table1->Active = false;
  Table1->DatabaseName = cbAlias->Text;
  Table1->TableName = cbTable->Text;
  Table1->Active = true;
  if(Table1->Active) Table1->GetFieldNames(cbField->Items);
}
//———— Button of Import —————
void __fastcall TForm1::Button1Click(TObject *Sender)
{ AnsiString Info;
  if(Table1->Active)
  { int t=Table1->FieldCount;
    for(i=3; i<Rec_Num+3; i++) // str_num
    { Table1->Append();
      for(j = 1; j < t; j++) // col_num
      { Label3->Caption=t;
        Table1->FieldValues["OBJNUM"] = Ob_Num;
        Info = Table1->Fields->Fields[j]->FieldName;
        Label1->Caption=j;
        Excel->SetSelection(i,j,i,j); // choice of Range
        if(Excel->Text != 0)
        { Label2->Caption=Excel->Text;
          Table1->FieldValues[Info] = Excel->Text; }
      } // for j
      Table1->Post();
    } // for i
  } // for if()
}
```



ნახ.2.11

2.7. OLAP კუბში მონაცემთა აგრეგაცია

ნაშრომში მოცემულია მონაცემთა ფორმალიზების აგრეგაციის მექანიზმი, სადაც ფართოდ გამოიყენება გრაფების თეორია.

მონაცემთა მრავალგანზომილებიანი მოდელის ძირითადი პრინციპებია:

მაჩვენებელიანი – სიდიდე (იგი ჩვეულებრივი რიცხვია), რომელიც გამოიყენება საგნობრივი ანალიზისათვის. მაგ რომელიმე პროდუქტის ყიდვა-გაყიდვის მოცულობა, ან ამ პროცესიდან მიღებული შემოსავალი. ერთი OLAP- კუბი შეიძლება შეიცავდეს ერთ ან რამდენიმე მაჩვენებელს.

განზომილება (dimension) – ერთი ან რამდენიმე სახის ობიექტთა სიმრავლეა, რომელიც ორგანიზებულია იერარქიული სტრუქტურით

და უზრუნველყოფს რიცხვითი მაჩვენებლის საინფორმაციო კონტექსტს. გაზომვა ვიზუალურად უნდა აისახებოდეს მრავალგანზომილებიანი კუბის გვერდებზე.

ობიექტები – მათი ერთობლიობა წარმოქმნის განზომილებას, ანუ ევრეთწოდებულ განზომილებათა წევრებს (members). განზომილებათა წევრები ვიზუალურად აისახება, როგორც ცალკეული წერტილები ან უბნები, რომელიც მოთავსებულია ჰიპერკუბის ღერძებზე. მაგ., დროითი განზომილება: დღე, თვე, კვარტალი, წელი: 26 მაისი 2002 წელი, მე-2 კვარტალი 2002 წელი და 2002 წელი. განზომილებაში მოცემული ობიექტები შეიძლება იყოს სხვადასხვა სახის მაგ: „მწარმოებელი“, „პროდუქტის საფირმო ნიშანი“, „წელი“, „კვარტალი“. ეს ობიექტები ორგანიზებული უნდა იყოს იერარქიული სტრუქტურით. ერთი სახის ობიექტებს შეესაბამება, იერარქიის მხოლოდ ერთი დონე.

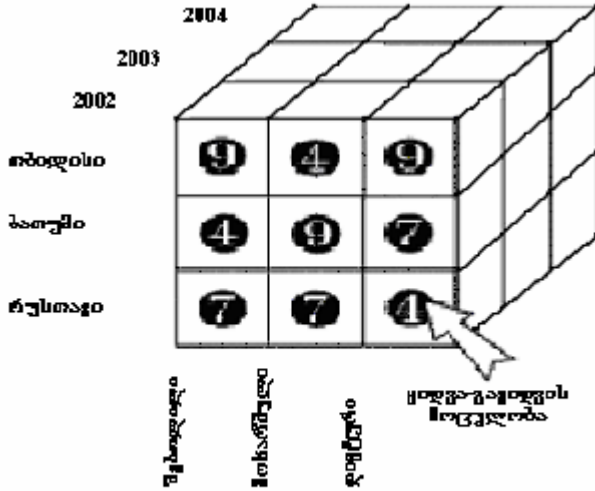
უჯრედი (cell) – კუბის ატომური სტრუქტურაა, რომელიც შეესაბამება რომელიმე მაჩვენებლის კონკრეტულ მნიშვნელობას. უჯრედები ვიზუალურად მოთავსებულია კუბის შიგნით და მათზე ასახულია მაჩვენებლების შესაბამისი მნიშვნელობები.

განზომილება თამაშობს ინდექსების როლს, იგი გამოიყენება ჰიპერკუბის უჯრედში მაჩვენებლის მნიშვნელობის იდენტიფიკაციისათვის. სხვადასხვა განზომილებით მიღებული კომბინაციები თამაშობს კოორდინატების როლს, რომელიც განსაზღვრავს კონკრეტული მაჩვენებლის მნიშვნელობას.

ყველა განზომილების შედეგად მიღებული კომბინაციებით შესაძლებელია რამდენიმე უჯრედის მნიშვნელობის განსაზღვრა. ამიტომ უჯრედის ერთმნიშვნელოვანი იდენტიფიკაციისათვის უცილებელია მიუთითოთ განზომილების ყველა კომბინაცია და მაჩვენებელი. ამავე დროს დაცული უნდა იყოს განზომილების იერარქიული დონეები, რათა აგრეგაციისა და მაჩვენებელთა მნიშვნელობის დეტალიზაციისათვის ამას დიდი მნიშვნელობა ენიჭება.

არსებობს იერარქიის შემდეგი სახე: **ბალანსირებული (balanced)** – იერარქია, რომელშიც დონეთა რიცხვი განისაზღვრება მისი სტრუქტურით და მუდმივია ამავე დროს, იერარქიული ხის თითოეული ტოტი შეიცავს ობიექტებს ყოველი დონიდან. მაგ.

ქევის ნებისმიერ მწარმოებელს შეუძლია აწარმოოს რამდენიმე სახის ქევი სხვადასხვა საფირმო ნიშნით (ნახ.2.12).



ნახ.2.12 სამგანზომილებიანი კუბი

ამ შემთხვევაში შეგვიძლია ვისაუბროთ სამდონიან იერარქიაზე, რომლის პირველ დონეზე იქნება ‘მწარმოებელი’, მეორეზე ‘სახეობა’ ხოლო მესამეზე ‘საფირმო ნიშანი’.

ბალანსირებული იერარქიის ფორმირებასათვის აუცილებელია არსებობდეს კავშირი „ერთი-მრავალთან“. მისი ყოველი დონე შეიძლება წარმოვადგინოთ, როგორც ცალკეული განზომილება, მაგრამ ეს განზომილება დამოკიდებული იქნება სხვა რომელიმე განზომილებაზე, ამიტომ საჭირო გახდება კუბის გაკვეთა.

დაუბალანსებელი (unbalanced) – იერარქია რომელშიც დონეთა რიცხვი შესაძლებელია შევცვალოთ, ხოლო იერარქიული ხის თითოეული ტოტი შეიძლება შეიცავდეს ობიექტებს არა ყოველი დონიდან, არამედ მხოლოდ პირველიდან. აუცილებელია აღვნიშნოთ, რომ დაუბალანსებელ იერარქიაში არსებული ყველა ობიექტი არის ერთი სახის. მაგ., იერარქია „ხელმძღვანელი–დამფასებელი“ ყველა ობიექტი ერთიანდება სახეში: „თანამშრომელი“.

არათანაბარი იერარქია, რომელშიც დონეთა რიცხვი განისაზღვრება მისი სტრუქტურით და მუდმივია.

ბალანსირებულისგან განსხვავდება მხოლოდ იმით, რომ იერარქიული ხის რომელიმე ტოტი შეიძლება არ შეიცავდეს რომელიმე დონის ობიექტებს. ასეთი სახის იერარქია შეიცავს ისეთ წევრებს როგორცაა ე.წ. ლოგიკური „მშობელი“ და იგი უშუალოდ არ მდებარეობს რომელიმე დონეზე.

ტიპიურ მაგალითს წარმოადგენს გრაფიკული იერარქია, რომელსაც შემდეგი დონეები აქვს: „ქვეყანა“, „შტატი“ და „ქალაქი“. თუ მონაცემთა ნაკრებში მოცემული გვექნება ქვეყნები, რომელთაც არ გააჩნიათ „შტატი“, მაშინ იგი მოთავსდება დონეებს შორის „ქვეყანა“ და „ქალაქი“.

ჩვენ განვიხილავთ მხოლოდ **ბალანსირებულ** იერარქიას, რადგან **დაუბალანსებელი** და **არათანაბარი** პრაქტიკაში თითქმის არ გამოიყენება.

აგრეგატებს უწოდებენ განსაზღვრული პირობის შესაბამისად მახასიათებლის საწყისი მნიშვნელობათა აგრეგირებას. აქ იგულისხმება მცირე რაოდენობის აგრეგატების მნიშვნელობათა ფორმირების ნებისმიერი პროცედურა, რომელიც ეყრდნობა დიდი რაოდენობის საწყის მნიშვნელობებს.

ნაწილობრივი და სრული აგრეგაცია

კუბის აგრეგაციის ხარისხი გამოითვლება ფორმულით:

$$\alpha = \frac{a}{a^*}$$

სადაც a —არის აგრეგირებულ მახასიათებლის მნიშვნელობის რეალური რაოდენობა, a^* - კუბის საწყისი მონაცემების აგრეგატული მნიშვნელობის მაქსიმალური რაოდენობა [31].

ფორმულის გამოყენებისას a და a^* - სთვის განვიხილოთ მარტივი შემთხვევები, ორ და სამ განზომილებიანი სივრცისთვის, რასაც მოგვიანებით გაერთიანებული სახით ვნახავთ. იგივე შეეხება განზომილებებში დონეთა იერარქიას: თავიდან განიხილება მარტივი შემთხვევა (ერთ დონიანი), ხოლო შემდგომ განვიხილავთ რამოდენიმე დონეს, რომელიც ფორმულაში წარმოდგენილი იქნება გაერთიანებული სახით. ასეთი მიდგომა აიოლებს მახასიათებლის აგრეგატული მნიშვნელობის მიღების პროცესს და თვალსაჩინოს ხდის დიდი განზომილების სივრცეების წარმოდგენის მექანიზმს.

გამოყენებული ლიტერატურა:

1. ბოტჭე კ., სურგულაძე გ., დოლიძე თ., შონია ო., სურგულაძე გ. თანამედროვე პროგრამული პლატფორმები და ენები // თბილისი, 2003.
2. გოვინიაშვილი გ., სურგულაძე გ., შონია ო. დაპროგრამების მეთოდები // თბილისი, 1997.
3. თურქია ე., ბიზნეს-პროცესების მართვის მათემატიკური მოდელის დამუშავება პროცეს-ორიენტირებული მიდგომით // სტუ-ს შრომები, № 1(451), თბილისი, 2004.
4. თურქია ე., სურგულაძე გ. ბიზნეს-პროცესების ნაკადების მართვის ინტერნეტული მხარდაჭერი სისტემა XML და Java ენების საფუძველზე // ჟურნალი „ინტელექტი“, № 2(16), თბილისი, 2003.
5. იაბლონსკი შ., სურგულაძე გ., თურქია ე., ჩიხრაძე ბ. ბიზნეს-პროცესების ნაკადთა მართვის ავტომატიზებული სისტემა // სტუ-ს შრომები, № 4(437), თბილისი, 2001.
6. კოტლერი ფ., მარკეტინგის საფუძვლები // თარგმ. ინგლ., თბილისი, 1993.
7. რეისიგი ვ., სურგულაძე გ., გულუა დ., ვიზუალური ობიექტ-ორიენტირებული დაპროგრამების მეთოდები // თბილისი, 2002.
8. სურგულაძე გ., თურქია ე. ბიზნეს-გეგმის ავტომატიზებული დამუშავების პროცესის სამუშაო ნაკადების მართვის სისტემა // სტუ-ს შრომები, № 8(424), თბილისი, 1998.
9. სურგულაძე გ., დოლიძე თ., თურქია ე., ოხანაშვილი მ. საწარმოო ფირმებში მარკეტინგული პროცესების მართვის ინფორმაციული სისტემის დაპროექტება და რეალიზაცია UML-ტექნოლოგიით //სტუ-ს შრომები, № 4(437), თბილისი, 2001.
10. სურგულაძე გ., თურქია ე. ბიზნეს-პროცესების მართვის ავტომატიზებული სისტემების დაპროექტება // მონოგრაფია, სტუ, თბილისი, 2003.
11. სურგულაძე გ., ქაშიბაძე მ. ოპერაციული სისტემები // ლექციების კონსპექტი, თბილისი, 1993.
12. ჩოგოვაძე გ., ინფორმაცია // თბილისი , 2003
13. Barker R. CASE*Method. Entity-Relationship Modelling. Copyright Oracle Corporation UK Limited, Addison-Wesley Publishing Co., 1990.
14. Bohm M., Entwicklung von Workflow-Typen//Springer Verlag, Heidelberg, 2000
15. Brackett J., C. McGowan. Applying SADT to Large System Problems // SofTechTechnical Paper TP059,January 1977.
16. Business Process Re-engineering: The Oracle Perspective. ORACLE CONSULTING, 1994

17. Bußler C., Organisationsverwaltung in Workflow-Management Systemen.//Deutscher Universitätsverlag, Germany, 1998.
18. Connor M. Structured Analysis and Design Technique - SADT.//Auerbach portfolio 32-04-02, 1979.
19. Case Method- <http://www.interface.ru/fset.asp?Url=/case/sadt4.htm>
20. David Hay, XML: What is It, Anyway? www.essentialstrategies.com
21. Document Object Model (DOM) Level 1 Specification <http://www.w3.org/TR/1998/REC-DOM-Level-1-19981001/>.
22. e-commerce-
http://www.microsoft.com/rus/sql/casestudies/ecommerce_site.asp
23. Gierhake O., Integriertes Geschäftsprozessmanagement//Vieweg Verlag, Gemany, 1999.
24. Griese J., Sieber P., Betriebliche Geschäftsprozesse// Germany, 2001.
25. Hammer M., Champy J. Reengineering the Corporation. A Manifesto for Business Revolutions. Harper Business, 1993.
- 26.Harrington J. Business Process Improvement. New York: McGrawHill, 1991.
27. Hickman L.J. Technology and Business Process Reengineering (BPR): Understanding Where BPR Fits into the World of Information Systems Developers. Proc. of ORACLE User Forum 93, v.2, Vienna, 1993.
28. Hunter J. McLaughlin B. Easy Java/XML integration with JDOM, Part2 //http:// www.javaworld.com /javaworld, 2000.
29. Jablonski St., Prozessinnovation durch Workflow – Management – Systeme // Germany, 1995.
30. Jablonski St., Workflow – Management – Systeme // Germany, 1995.
- 31.Jablonski St., Architektur von Workflow – Management – Systemen // Informatik forschung und Entwicklung, Germany, 1997.
32. Jablonski St., Von der Anwendungsanalyse zu ersten Systemkonzepte für Workflow – Management – Lösungen// Arbeitsbereich des Institut Informatik der Universität Erlangen – Nürnberg, Germany, 2000.
33. Jacobson I., Ericsson M., Jacobson A. The Object Advantage: Business Process Reengineering with Object Technology. ACM Press. – Addison – Wesley Publishing, N – Y: 1995.
34. Java:http://www.galileocomputing.de /openbook/javainse13/javainse1_000000.htm #Xxx998140 20.1
35. Kaldich R., Von Prozessmonitoring zum Prozessmanagement //Europäischer Verlag der Wissenschaften, Bern, 1998.

36. Lehmann R., Fachlicher Entwurf von Workflow – Management – Anwendungen // Europäischer Verlag der Wissenschaften, Bern, 1998.
37. Martens A., Software – Engineering von Workflow – Applikationen mit Petrinetzen // Germany, 1999.
38. Meffert H., Marketing, Grundlagen der Absatzpolitik; mit Fallstudien Einfuhrung und Relaunch des VW – Golf // H.Meffert. – 7., uuberarb. u. erw. Aufl., Nachdr. – Wiesbaden: Gabler, 1993.
39. Merz M., Electronic commerce: Marktmodelle, Anwendungen und Technologien // dpunkt – Verlag, Germany, 1999.
40. Neeb J., Administration of Workflow Management Solutions // Workshop, Philaelpia, 2000.
41. Otey M., Conte P. SQL Server 2000, Developer's Guide. Osborn, 2001.
42. Rauffer H., Dokumentorientierte Modellierung und Controlling von Geschäftsprozessen. Deutscher Universitätsverlag, Gemany, 1997.
43. Reisig W., Petrinetze – eine Einführung // Germany, 1982.
44. Rein G., Organization Design Viewed as a Group Process Using Coordination Technology // MCC Technical Report Number CT – 039 – 92, Germany, 1991.
45. Sceppa D. Microsoft ADO.NET, Microsoft Corporation, 2002.
46. Schamburger R., Integrierte Betrachtung von Anwendung und Systemen zur verteilten Workflow // Germany, 2000.
47. Schuster H., Architektur verteilter Workflow – Management – Systeme // Germany, 1997.
48. Sharp A., McDermott P., Workflow Modeling: Tools for Process Improvement and Application Development // Norwood, 2001.
49. Sheer A., ARIS – Modellierungsmethoden, Metamodelle, Anwendungen. // Berlin, 1998.
50. Stein K., Integration von Anwendungsprozess – modellierung und Workflow – Management // Disertation, Germany, 1999.
51. Wedekind H., Programming in the large und der Entwurf Workflow – Schemata // Unveröffentlichtes Manuskript, Erlangen, Germany, 1996.
52. Андерсен Б. М. Бизнес – процессы: Инструменты совершенствования. (пер. с англ. Я. Ариничесвая С. В.) // Москва 2000.
53. Я. XSL: путь к элегантности // www.w3.org/Tr/WD-xsl/
54. Электронный бизнес // <http://www.btime.ru/econs/biz/1.htm>
55. Электронная торговля без проблем // <http://www.w3.org/Tr/WD-xsl/>
56. Использование ЯББ // ХМ – обработчик // www.mrcpk.nstu.ru/km/

- 57) Кобайлас Дж. Network Work. Ежедневник "Computerworld" и Ньюбик США 1998.
- 58) Мейнджера Дж. Основы программирования: Учебник Киев 1996
- 59) Прикладные программные интерфейсы и архитектура Учебник для ХМ
- [Итp:ььwwwььfvььeduььcyberььCEььCEььImььNsььEctiььPO=41&&BrticььM=10](http://www.fv.edu/cyber/CE/CE/Im/secti/PO=41&&BrticM=10)
- 60) Роббинс Р. Коултер М. Менеджменты Москв 2000
- 61) Четвергическое планирование корпоративных сетей
[ИтP:ььwььBььMььmsььxььuььCььEььtsььBrticььNsььььBrt_ььtььxььImьь](http://www.fv.edu/cyber/CE/CE/Im/secti/PO=41&&BrticM=10)
- 62) Федорова Б. Менеджмент внедрения информационных технологий в систему управления предприятия "Корпоративный Менеджмент"
- 63) Шеер А. Бизнес процессы. Основные понятия. Теория. Методы: Учебник ВЕК РИ №4 Издательство Перхьянгл 1999
- 64) Шелли П. Чего мы ждем от ХМ? "Мир ПК" №9 1998
- 65) Черченя У. Аюфья Р. Арнофья Л. Введение в исследование операций Москв 1968
- 66) Язык ХМ — язык творчества (www.xim.ru)