

MANAGEMENTUL RISCULUI PRIN ANALIZA MULTIDIMENSIONALA A DATELOR

Constanța Zoie Rădulescu

Institutul Național de Cercetare - Dezvoltare în Informatică

Rezumat: Lucrarea prezintă un model multidimensional de organizare a datelor și o analiză de tip OLAP (prelucrare analitică on line) pentru managementul riscului de poluare a apei în industria textilă. Se descriu, mai întâi, etapele necesare a fi rezolvate într-un proces de management al riscului de poluare a apei. Se prezintă modelul multidimensional al datelor: dimensiunile, faptele și măsurile folosite precum și legăturile dintre acestea într-o structură de tip stea. Se exemplifică modelul pentru cinci poluanți caracteristici industriei textile. Se stabilesc ierarhiile dimensiunilor și se definesc operațiile tipice OLAP ce se pot aplica modelului multidimensional al datelor. Se exemplifică operația de feliere și decupare de cubulețe (slice and dice). Se definesc apoi măsurile teoretice și practice pentru calculul riscului de poluare. În final, se calculează riscul pentru poluanții luați în considerare.

Cuvinte cheie: prelucrare analitică on line (OLAP), model multidimensional, operații OLAP, risc, managementul riscului.

1. Introducere

Unitățile economice, din ramura industriei textile, care produc țesături, folosesc în procesul tehnologic un număr mare de materii prime textile, substanțe chimice, produse auxiliare și coloranți. Acestea, alături de folosirea unor tehnologii și utilaje uzate moral, conduc la poluarea apei, aerului și solului. Apele reziduale, rezultate în procesele de prelucrare chimică, sunt caracterizate de o mare varietate de factori poluanți. De multe ori, aceste ape trebuie tratate prin stații de epurare, înainte de a fi deversate în canalizarea orașelor sau în apele curgătoare.

Impactul pe care aceste ape industriale insuficient purificate, le au asupra mediului este semnificativ prin schimbarea în timp a echilibrului ecologic. De aceea, controlul poluării mediului, încadrarea calității acestuia în norme și standarde impuse, managementul riscului de poluare sunt unele dintre problemele cu care se confruntă managerii acestor unități economice.

Pe de altă parte, se dorește creșterea profitului unităților economice, a eficienței economice și o dezvoltare economică. Această dezvoltare trebuie, însă, să fie strâns legată de impactul pe care îl are asupra mediului, pentru că numai astfel se produce o dezvoltare economică durabilă. Pentru a fi durabilă, dezvoltarea economică trebuie să îmbine armonios eficiența economică cu protecția mediului și cu restaurarea sistemelor ecologice.

Pentru a rezolva cele două probleme ale căror obiective sunt conflictuale, reducerea poluării și creșterea eficienței economice, managerii unităților economice trebuie să găsească o soluție de compromis. Instrumentele software care îi pot ajuta pe aceștia să găsească o soluție și să ia cele mai bune decizii sunt SSD (sistemele suport de decizie). Marea majoritate a SSD destinate rezolvării acestor tipuri de probleme sunt bazate pe metode de analiză multicriterială și managementul riscului. O organizare a datelor într-un model multidimensional și o analiză de tip OLAP permite managerilor evidențierea esenței datelor, analiza interactivă și rapidă a acestora.

Scopul lucrării este de a prezenta un mod de organizare a datelor într-un model multidimensional și o analiză OLAP pentru managementul riscului de poluare a apei în industria textilă.

Lucrarea este organizată după cum urmează. În capitolul 2, se descriu etapele necesare a fi rezolvate într-un proces de management al riscului de poluare a apei. Capitolul 3 prezintă modelul multidimensional al datelor propus, iar capitolul 4 prezintă analiza modelului multidimensional prin operații OLAP. În capitolul 5, se definesc câteva măsuri pentru calculul riscului, iar în capitolul 6, se prezintă exemplul numeric pentru calculul riscului.

Lucrarea se încheie cu concluzii și cu o bibliografie.

2. Etape în procesul de management al riscului de poluare a apei

Etapele necesare în procesul de management al riscului de poluare a apei sunt:

1. preluarea datelor din bazele de date operaționale și crearea unei colecții de date specifice datamart;
2. organizarea datelor într-un model multidimensional;
3. analiza modelului dimensional prin operații OLAP;
4. definirea de măsuri pentru calculul riscului;
5. calculul riscului prin măsuri specifice și analiza rezultatelor obținute.

Vom explica detaliat un exemplu numeric pentru etapele 2-5 ale acestui proces.

Date privind măsurarea poluării apei industriale prin preluarea și analizarea de probe, la diverse momente de timp și pentru diverși poluanți (caracteristici industriei textile) există în baze de date operaționale, la nivelul unităților economice, precum și la unități specializate în preluarea de probe și analizarea lor. Cu ajutorul acestor date se poate crea o colecție de date specifice – datamart - pentru monitorizarea riscului de poluare a apei. Aceste colecții de date specifice, numite datamarts, sunt destinate managerilor, analiștilor și specialiștilor angrenați în luarea deciziilor strategice privind dezvoltarea și viitorul unităților economice.

Instrumentele specializate pentru asistarea deciziilor, care transformă informațiile din datamart în forma cerută de decidenți și oferă posibilitatea analizei tendințelor și corelațiilor sunt *prelucrarea analitică – on line* („*On line analytic processing*” – OLAP) și *mineritul datelor* („*Data mining*” – DM).

Conform cu glosarul propus de OLAP Council, [3] și [2], „prelucrarea analitică on-line (OLAP) desemnează o categorie de instrumente software care permit analiștilor, managerilor și directorilor să înțeleagă esența datelor printr-un acces rapid, consistent și interactiv, la o mare varietate de vederi posibile ale informațiilor care au fost obținute prin transformarea datelor primare, astfel încât să reflecte dimensiunile reale ale întreprinderii așa cum o înțelege utilizatorul”.

Instrumentul OLAP se bazează pe modelarea multidimensională a datelor și permite analiza interactivă și rapidă a datelor prin operații de tip rulare (roll-up), forare (drill-down), feliere și decupare de cubulețe (slice and dice) etc. Utilizatorii pot obține rezultate imediate prin formularea de întrebări, parcurgând dinamic dimensiunile structurii multidimensionale, lucrând cu niveluri diferite de sinteză sau detaliere.

Mai multe informații privind OLAP se găsesc în [1-6].

3. Modelul multidimensional al datelor

Modelul multidimensional permite modelarea și vizualizarea datelor în dimensiuni multiple.

El este definit prin dimensiuni și fapte. „Dimensiunile” conțin descrierea informațiilor care dau înțeles numerelor conținute în tabelul „fapte”. „Dimensiunile” conțin, de obicei, valori alfanumerice. „Fapte” conțin valori numerice, pe care dorim să le analizăm. Dimensiunile și faptele sunt legate între ele prin structuri de tip stea, fulg de zăpadă sau constelații.

„Dimensiunile” luate în considerare pentru exemplul nostru sunt:

1. „timpul” (zi, lună, an),
2. indicatorii ce descriu concentrația poluanților în apă: această dimensiune o vom numi generic „poluanți”,
3. probele de apă ce sunt preluate și analizate la diverse momente de timp, în diverse locuri de preluare, pentru diferite surse de poluare. Pe baza acestora se măsoară diferiți indicatori ai concentrației poluanților în apă. Această dimensiune o vom numi generic „probe poluare”,
4. unitățile economice.

Deci, mulțimea dimensiunilor este {timp, poluant, probe poluare, unitate economică}.

Aceste dimensiuni permit un management al riscului de poluare a apei.

Fiecare „dimensiune” are un tabel asociat, numit tabel dimensiune, care descrie atributele dimensiunii. De exemplu, un tabel dimensiune pentru poluant poate conține atributele: cod poluant, nume poluant, unitatea de măsură, pragul de alertă, concentrația maxim admisă, norma de mediu, care stabilește valorile pentru pragul de alertă și pentru concentrația maxim admisă. Un exemplu pentru apă industrială, tratată prin stația de epurare, este prezentat în tabelul 1.

Tabelul 1. Dimensiunea poluant

Cod	Indicatori de poluare	U.M.	Norma	Prag de alertă	Concentrație maxim admisă
1.	Materii în suspensie	mg/l	Normativul NTPA 002	210	300
2.	Consum biochimic de oxigen la 5 zile (CBO ₅)	mg O ₂ /l	Normativul NTPA 002	210	300
3.	Consum chimic de oxigen - met. cu bicromat de potasiu (CCO-Cr)	mg O ₂ /l	Normativul NTPA 002	350	500
4.	Substanțe extractibile cu eter de petrol	mg/l	Normativul NTPA 002	14	20
5.	Detergenți sintetici anion-activi biodegradabili	mg/l	Normativul NTPA 002	21	30

Pentru exemplul nostru vom considera „tabelul dimensiune poluant” simplificat, având doar cod poluant și nume poluant.

„Fapte” au, de asemenea, un tabel asociat, reprezentat în exemplul nostru de tabelul „Analiza riscului”. Faptele sunt valori numerice care reprezintă subiectul unei analize. Faptele sunt implicit definite prin combinația elementelor dimensiunilor. Un fapt (celulă) are o anumită granularitate (grad de detaliere) determinată de combinația dimensiunilor. În exemplul nostru, o granularitate este valoarea maximă a concentrației unui poluant măsurată într-o probă, la un moment dat în timp.

Faptele cuprind și măsurile pentru care se va realiza analiza. O măsură are două componente: proprietatea numerică a faptului, de exemplu valoarea măsurată a concentrației unui poluant în apă și o formulă (adesea o funcție de agregare simplă cum ar fi max, min, sumarea sau numărarea), de exemplu maximul valorii măsurate a concentrației unui poluant în apă.

Măsurile considerate pentru problema de analiză a riscului, sunt:

1. valoarea maximă măsurată,
2. numărul de valori ce depășesc pragul de alertă stabilit,
3. numărul de valori ce depășesc concentrația maximă permisă pentru un poluant sau categorie de poluanți și
4. numărul valorilor ce au fost măsurate.

Măsurile pot lua diferite valori pentru combinații diferite ale valorilor dimensiunilor.

Este important a distinge între măsuri care se pot aplica pentru toate dimensiunile (cum ar fi numărul de valori ce depășesc pragul de alertă stabilit, numărul de valori ce depășesc concentrația maximă permisă pentru un poluant și numărul valorilor ce au fost măsurate) și măsuri care au sens numai pentru o dimensiune (de exemplu valoarea maximă măsurată are sens doar pentru dimensiunea poluant).

Tabelul „Fapte” conține, de asemenea, coduri de legătură cu tabelele dimensiune.

Fiecare tabel dimensiune conține un cod de legătură cu tabelul fapte.

Măsurătorile poluării apei industriale s-au obținut prin analiza probelor preluate, pentru diferite momente de timp, pentru diverși indicatori și probe de analiză a poluării apei.

Legătura dintre tabelele dimensiune și tabela fapte este o structură de tip stea. Structura stea este cel mai comun model de date. Ea conține tabelul central, care este tabelul fapte (cu volum mare de date) și o mulțime de tabele dimensiune (în cazul nostru patru), dispuse împrejurul lui.

Acest model de tip stea este prezentat în figura 1.

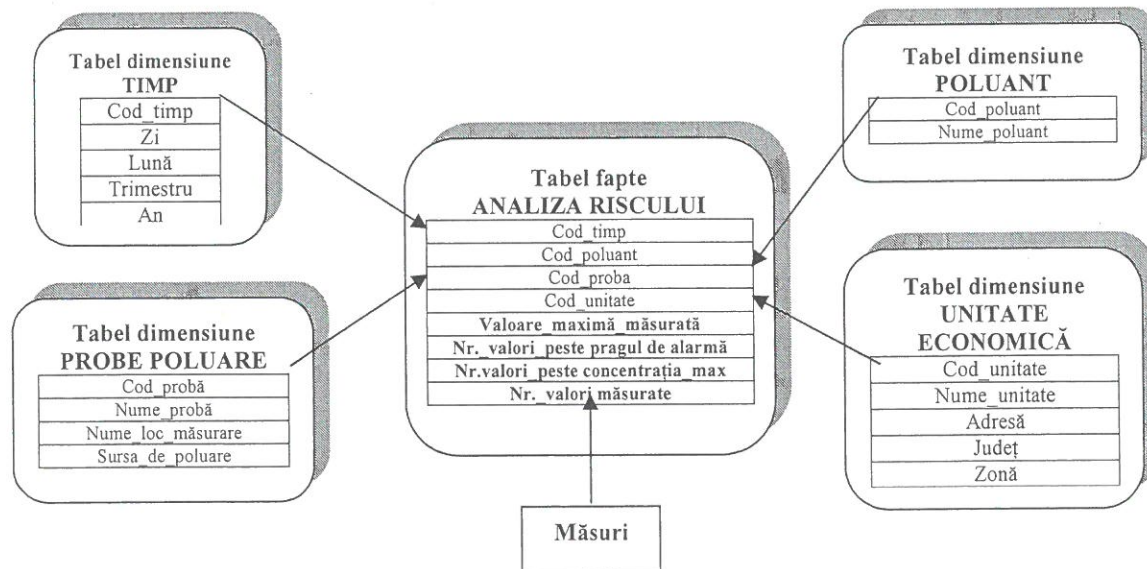


Figura 1. Modelul de tip stea pentru problema poluării apei

Schema de tip stea prezentată este suficient de generală pentru a putea fi aplicată și altor probleme de poluare a mediului (aer, sol), cât și pentru alte unități economice care poluează, din alte ramuri industriale diferite de cea textilă.

Schema de tip stea se poate transforma în schemă de tip fulg de nea atunci când se iau în considerare și alte dimensiuni ce se leagă de dimensiunile existente. Astfel, pentru exemplul nostru se pot considera tabele distincte pentru: a) norme (legi, hotărâri de guvern) privitoare la poluarea mediului și b) tipul de poluare (aer, apă, sol).

Modelul din exemplul nostru este 4 dimensional. Dacă luăm în considerare numai trei dintre dimensiuni și anume timp, probe poluare și poluant atunci modelul pentru exemplul nostru este prezentat în figura 2. Considerăm dimensiunea Poluant ca având 5 elemente. Acestea sunt:

1. Consum biochimic de oxigen la 5 zile (CBO₅),
2. Consum chimic de oxigen - metoda cu bicromat de potasiu (CCO-Cr),
3. Detergenți sintetici anion-activi biodegradabili,
4. Materii în suspensie,
5. Substanțe extractibile cu eter de petrol.

Considerăm dimensiunea „probe poluare” ca având 3 elemente, {proba 1, proba 2, proba 3}, iar dimensiunea „timp” ca având 4 elemente {L1, L2, L3, L4}, primele 4 luni ale anului 2004.

Numărul de celule ale modelului este egal cu 60 și reprezintă numărul maxim de înregistrări din tabelul fapte. Considerăm o singură măsură și anume valoarea maximă măsurată.

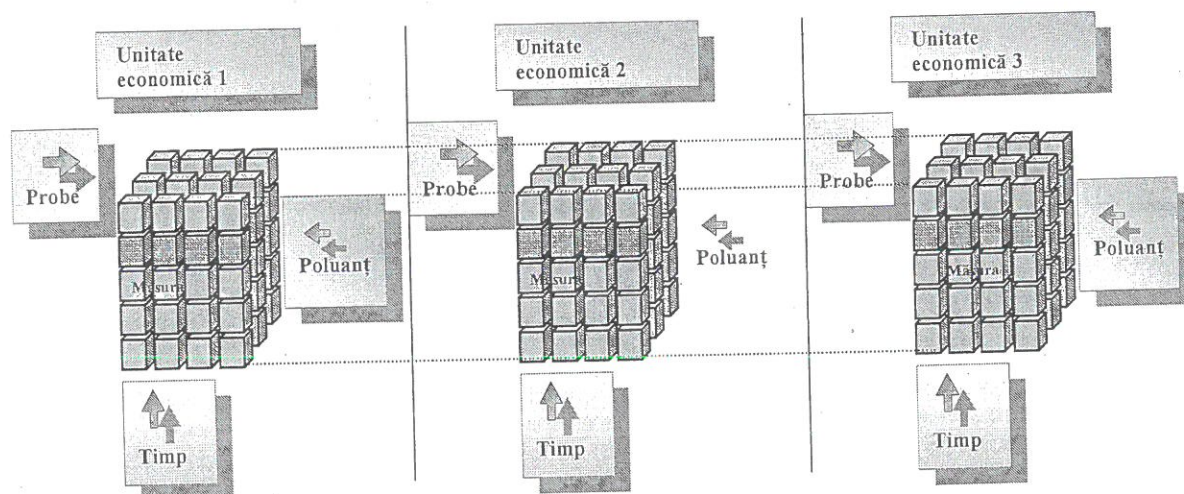


Figura 2. Organizare tridimensională a datelor privind poluarea apei în industria textilă

Atunci când se ia în considerare și a patra dimensiune și anume unitatea economică, organizarea datelor se poate reprezenta ca o mulțime de „cuburi” cu trei dimensiuni. Astfel, organizarea datelor pentru trei unități economice este prezentată în figura 3. Numărul de celule în acest caz este 180.

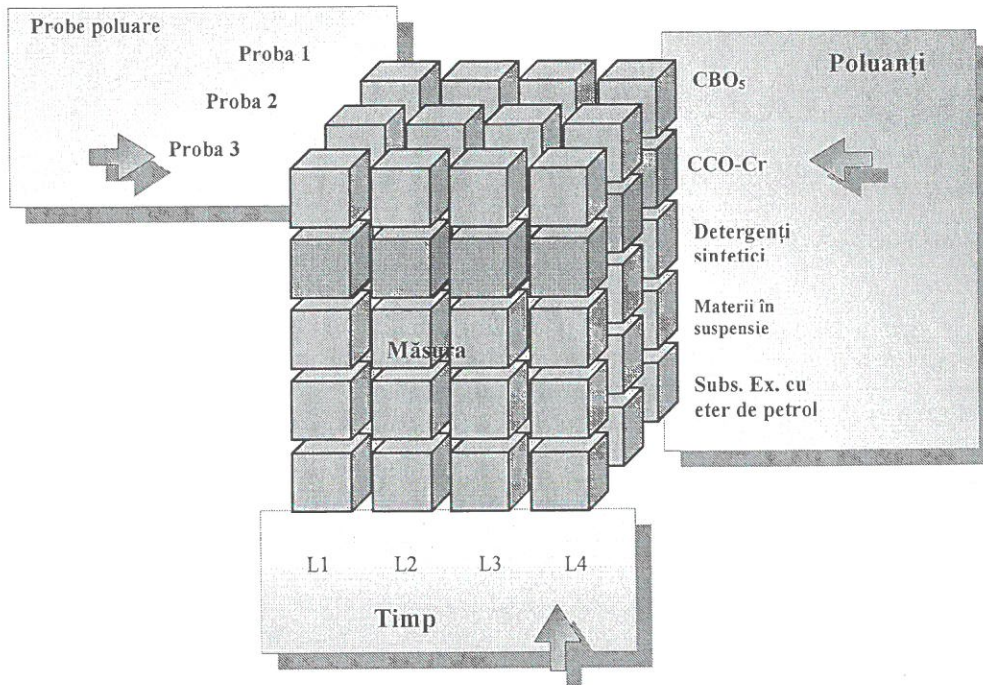


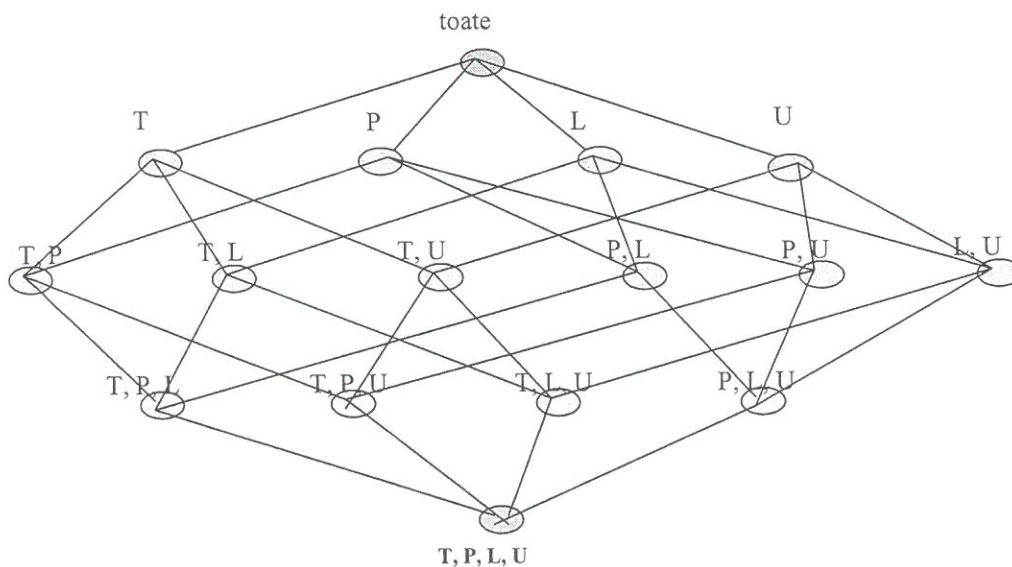
Figura 3. Organizarea datelor luând în considerare patru dimensiuni

Dacă se consideră mulțimea D cu elemente dimensiunile modelului n -dimensional $d_1, d_2, \dots, d_n : D = \{d_1, d_2, \dots, d_n\}$, atunci $P(D)$, mulțimea părților lui D este

$$P(D) = \{\emptyset, \{d_1\}, \{d_2\}, \dots, \{d_n\}, \{d_1, d_2\}, \dots, \{d_1, d_2, d_3\}, \dots, \{d_1, d_2, \dots, d_n\}\}, \text{ iar } |P(D)| = 2^n$$

Deci, sunt 2^n părți ale mulțimii D . Fiecare parte reprezintă o vizualizare (vedere) unică a datelor la un nivel dat de granularitate.

Selectând diverse părți din mulțimea D putem analiza datele în raport cu măsurile alese.



Pentru exemplul nostru mulțimea D are 4 elemente, iar mulțimea părților lui D , $|P(D)| = 2^4 = 16$.

Dacă notăm dimensiunile alese pentru exemplul nostru cu: T – timp, P – poluant, L – proba (loc de măsurare), U – unitate economică atunci reprezentarea acestor părți este prezentată în figura 4.

Primul nivel reprezintă cel mai înalt grad de sintetizare. Al doilea nivel reprezintă părțile cu o singură dimensiune, nivelul 3 - părțile cu două dimensiuni, nivelul 4 - părțile cu trei dimensiuni, iar nivelul 5 - cel mai mare grad de detaliere a datelor.

Dacă se consideră toate măsurile, un exemplu numeric pentru (T, P) este prezentat în tabelul 2. (T, P) este o parte a mulțimii D reprezentată de perechea timp și poluant. Se aplică toate măsurile definite pentru toate probele și toate unitățile economice. Aceste dimensiuni (probe poluare și unități economice) nu apar explicit.

De exemplu, la Consum biochimic de oxigen la 5 zile (CBO5), valoarea măsurată maximă pentru luna L1 pentru toate probele de poluare și toate unitățile economice este 363.6.

Tabelul 2. Exemplu pentru (T, P)

Unități economice	Toate	Luna			
Probe poluare	Toate	L1	L2	L3	L4
Poluant	Timp				
Consum biochimic de oxigen la 5 zile (CBO5)	Valoare măsurată maximă	363.6	344.6	350,6	363.6
	Nr. valori peste pragul de alertă	24	2	6	14
	Nr. valori peste concentrația maxim admisă	6	0	3	5
	Numărul de măsurători	25	7	37	27
Consum chimic de oxigen - metoda cu bicromat de potasiu (CCO-Cr)	Valoare măsurată maximă	547.1	542.4	541,6	546.1
	Nr. valori peste pragul de alertă	32	2	10	8
	Nr. valori peste concentrația maxim admisă	11	1	3	5
	Numărul de măsurători	33	7	27	35
Detergenți sintetici anion-activi biodegradabili	Valoare măsurată maximă	31	32	35	31
	Nr. valori peste pragul de alertă	1	7	7	2
	Nr. valori peste concentrația maxim admisă	1	1	4	1
	Numărul de măsurători	32	7	20	28
Materii în suspensie	Valoare măsurată maximă	394	399	310	370
	Nr. valori peste pragul de alertă	10	8	9	10
	Nr. valori peste concentrația maxim admisă	5	7	6	6
	Numărul de măsurători	32	19	23	34
Substanțe extractibile cu eter de petrol	Valoare măsurată maximă	45	40	50	40
	Nr. valori peste pragul de alertă	32	1	6	33
	Nr. valori peste concentrația maxim admisă	10	1	3	17
	Numărul de măsurători	33	7	25	40

4. Analiza modelului multidimensional prin operații OLAP

Modelul multidimensional permite obținerea de vizualizări ale datelor (existente și / sau calculate) pentru toate datele în cele patru dimensiuni precum și pe submulțimi ale acestora.

Un exemplu pentru vizualizări ale modelului multidimensional al datelor pentru 3 dimensiuni (poluant, probe poluare și timp) sunt prezentat în figura 5.

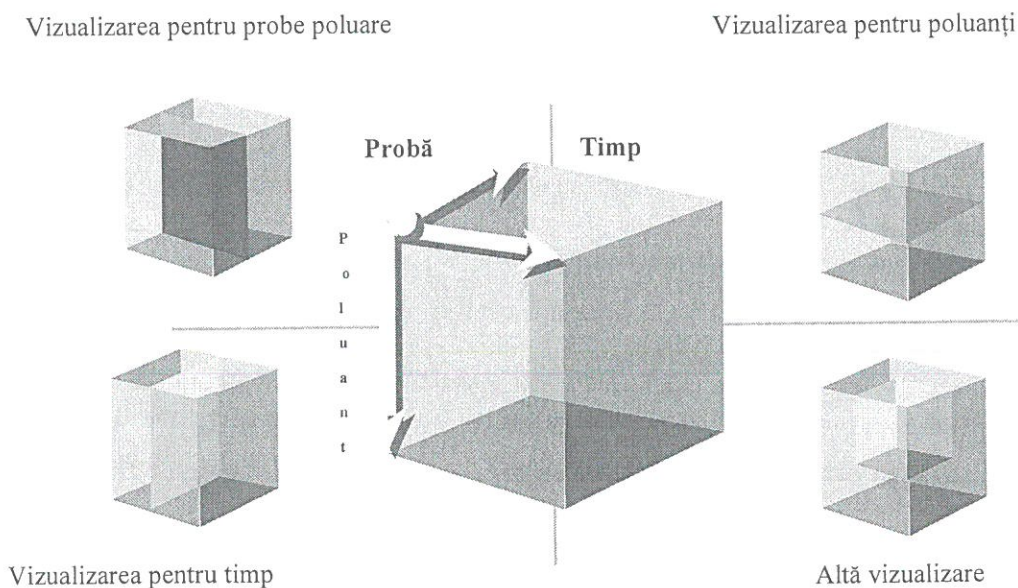


Figura 5. Vizualizări în modelul multidimensional 3 dimensional

Ierarhiile considerate pentru trei dimensiuni (unitate economică, probă poluare apă și timp) sunt prezentate în figura 6.

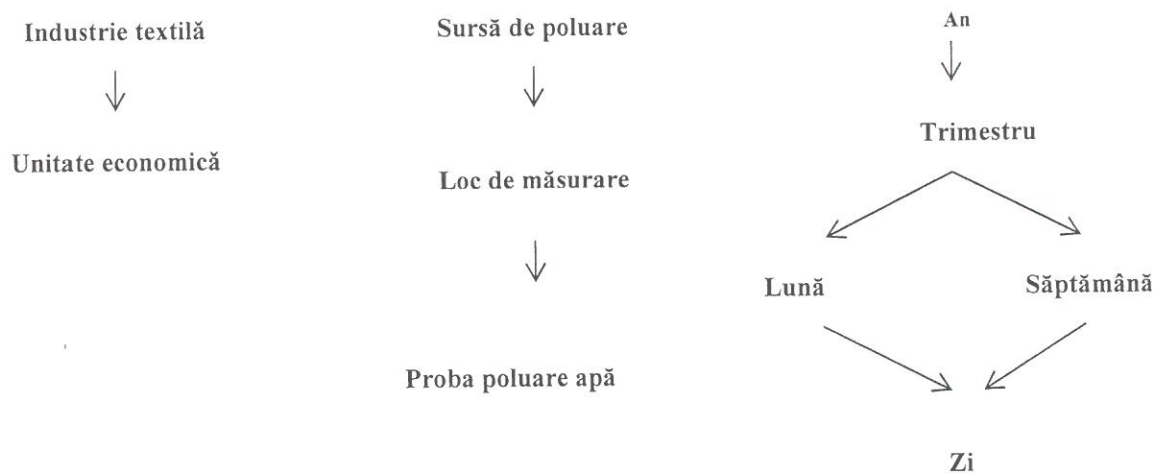


Figura 6. Ierarhiile pentru dimensiunile unitate economică, probe poluare și timp

Operațiile tipice OLAP asupra unui model multidimensional al datelor sunt rularea (roll-up), forarea (drill-down), felierea și decupare de cubulețe (slice and dice) și pivotarea (rotația).

Operațiile (interogările) OLAP au fost influențate de limbajul SQL și de calculul tabelar. O operație comună este de a agrega o măsură peste una sau mai multe dimensiuni. De exemplu:

- găsirea valorii maxime, măsurată pentru toată perioada de timp, pentru toate probele de măsurare a poluării, pentru poluantul materii în suspensie;
- găsirea numărului de valori măsurate ce depășesc concentrația maxim admisă, pentru trimestrul 4 anul 2004, pentru toți poluanții și toate probele de măsurare a poluării;

- găsirea numărului total de valori măsurate, pentru luna iunie anul 2004, pentru toți poluanții și probele de măsurare a poluării la intrarea în stația de epurare.

Operația de roll up (rulare) conduce la o sintetizare a datelor. Această sintetizare se realizează fie mergând de la un nivel inferior către un nivel superior într-o ierarhie a unei dimensiuni, fie prin reducerea dimensiunii.

Operația de drill down (forare) este inversă operației de roll up. Ea presupune trecerea de la un nivel superior de sinteză la un nivel inferior. Creșterea nivelului de detaliere la datelor se poate și realiza și prin adăugarea de noi dimensiuni.

Operațiile de roll up și drill down pot fi executate asupra componentelor unei ierarhii a unei dimensiuni. De exemplu, dacă aplicăm operația de drill down pentru un an, vom vedea descendenții direcți din ierarhia dimensiunii timp, adică trimestrele pentru an. Dacă aplicăm operația de roll up pentru trimestrul 1, pentru un an, vom ajunge la anul respectiv.

Operațiile de feliere și decupare de cubulețe (slice and dice) presupun:

- alegerea partiției pentru fiecare dimensiune a modelului multidimensional al datelor (ceea ce se realizează în interogări prin clauza "group by");
 - decuparea dintr-o partiție particulară de-a lungul uneia sau mai multor dimensiuni (corespunzător clauzei "where").
- Operația de pivotare presupune reorientarea cubului de date (3 D) pentru vizualizare în plane 2D.

Un exemplu de operație "slice and dice" este prezentată în tabelul 3. Se alege din cubul de date numai anul 2004, toate probele și 4 poluanți.

Tabelul 3. Operația de „slice and dice”

Unități economice	Toate	
Probe poluare	Toate	
Poluant	Timp	
Consum biochimic de oxigen la 5 zile (CBO5)	Valoare măsurată maximă	363.6
	Nr. valori peste pragul de alarmă	46
	Nr. valori peste concentrația maxim admisă	14
	Numărul de măsurători	96
Consum chimic de oxigen - metoda cu bicromat de potasiu (CCO-Cr)	Valoare măsurată maximă	547.1
	Nr. valori peste pragul de alarmă	52
	Nr. valori peste concentrația maxim admisa	20
	Numărul de măsurători	102
Detergenți sintetici anion-activi biodegradabili	Valoare măsurată maximă	35.00
	Nr. valori peste pragul de alarmă	17
	Nr. valori peste concentrația maxim admisă	7
	Numărul de măsurători	87
Substanțe extractibile cu eter de petrol	Valoare măsurată maximă	399
	Nr. valori peste pragul de alarmă	37
	Nr. valori peste concentrația maxim admisa	24
	Numărul de măsurători	108

5. Definirea de măsuri pentru calculul riscului

Dintre măsurile riscului, folosite pentru evaluarea gradului de poluare a apei amintim dispersia, semidispersia, probabilitatea de depășire a unui prag dat, deviația superioară față de un prag dat, deviația medie absolută, momentul parțial superior.

Una dintre cele mai cunoscute măsuri ale riscului de poluare este probabilitatea de depășire a unui prag dat pentru valoarea măsurată a concentrației de poluant.

Dacă notăm cu X variabila aleatoare care are ca semnificație concentrația poluantului și cu τ un prag de admisibilitate (prag de alarmă, concentrație maxim admisă etc.) pentru concentrația acestuia, atunci o măsură a riscului este:

$$R_{\tau}(X) = P(X \geq \tau) = 1 - F_X(\tau) \quad (1)$$

Am notat aici cu F_X funcția de repartiție asociată variabilei aleatoare X .

În practică, dacă avem niște măsurători asupra concentrației unui poluant x_1, x_2, \dots, x_n făcute la momentele t_1, t_2, \dots, t_n echidistante în timp, putem estima $R_\tau(X)$ cu:

$$R(\tau, x_1, x_2, \dots, x_n) = \frac{\text{card}\{i \in \{1, 2, \dots, n\} \mid x_i \geq \tau\}}{n}$$

O măsură mai complexă decât (1) este momentul parțial superior UPM (Upper Partial Moment) de ordinul α . Aici, parametrul α este un număr strict pozitiv. El se definește astfel:

$$UPM_\alpha(\tau, X) = E[(X - \tau)_+^\alpha] = \int_\tau^\infty (x - \tau)^\alpha dF_X(x) = \int_\tau^\infty (x - \tau)^\alpha f_X(x) dx \quad (2)$$

Am notat aici cu F_X funcția de repartiție a variabilei aleatoare X și cu f_X densitatea de probabilitate asociată variabilei aleatoare X .

În practică, dacă avem niște măsurători asupra concentrației unui poluant x_1, x_2, \dots, x_n făcute la momentele t_1, t_2, \dots, t_n echidistante în timp, putem estima $UPM_\alpha(\tau, X)$ prin:

$$UPM_\alpha(\tau, x_1, x_2, \dots, x_n) = \frac{\sum_{i \in A_\tau} (x_i - \tau)^\alpha}{n}$$

Am notat $A_\tau = \{i \in \{1, 2, \dots, n\} \mid x_i \geq \tau\}$.

Pentru $\alpha = 2$ obținem:

$$UPM_2(\tau, x_1, x_2, \dots, x_n) = \frac{\sum_{i \in A_\tau} (x_i - \tau)^2}{n}$$

Se observă că: $R_\tau(X) = \lim_{\alpha \downarrow 0} UPM_\alpha(\tau, X)$

Deci, prima măsură (1) poate fi considerată ca un caz particular al măsurii (2).

Managerii din domeniul mediului au aversiune mare față de riscul de poluare vor folosi UPM_α cu α luând valori mari. Cu cât sunt mai mari depășirile poluantului față de pragul τ , ele sunt amplificate de exponentul α . Astfel, managerul de mediu care folosește UPM_α va avea un risc mai mare cu cât α este mai mare.

6. Calculul riscului prin măsuri și analiza rezultatelor obținute

O analiză mai complexă asupra modelului multidimensional al datelor se realizează folosind măsuri ale riscului.

Vom folosi măsura (1) definită anterior, pentru cazul situațiilor practice. Pentru cei 5 poluanți valoarea riscului calculat este prezentată în tabelul 3.

Tabelul 3. Valorile riscului

Poluanți	Valoarea riscului
Consum biochimic de oxigen la 5 zile (CBO5)	0,1458
Consum chimic de oxigen - metoda cu CCO-Cr	0,2549
Detergenți sintetici anion-activi biodegradabili	0,1494
Materii în suspensie	0,2222
Substanțe extractibile cu eter de petrol	0,2952

Se constată că, pentru perioada de timp, poluanții și probele de analiză considerate, riscul cel mai mare de poluare a apei îl prezintă „Substanțele extractibile cu eter de petrol” (0,2952).

Grafic, valorile calculate pentru risc sunt prezentate în figura 7.

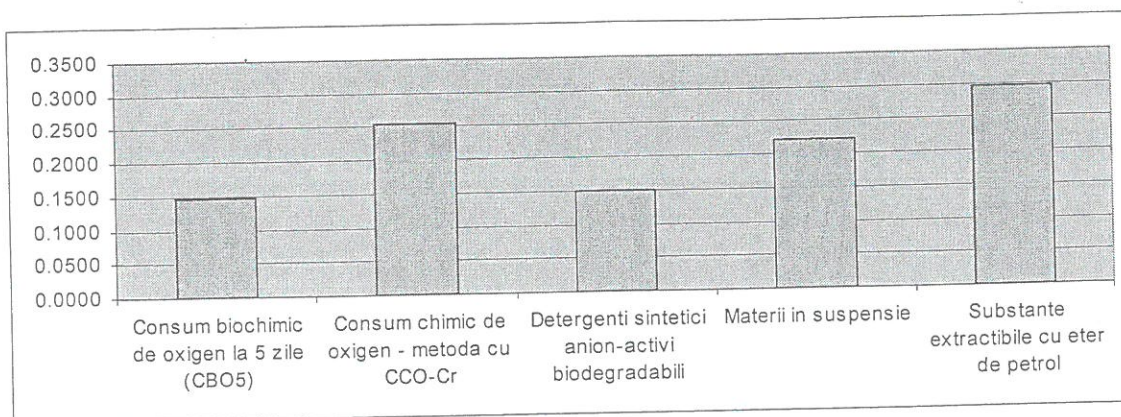


Figura 7. Graficul valorilor calculate ale riscului

7. Concluzii

Lucrarea a prezentat un model multidimensional de organizare a datelor, care permite analizarea colecției de date prin vizualizări ale unor cuburi multidimensionale. Mulțimea de vederi posibile asupra colecției de date atunci când se definesc măsuri convenabile poate fi de real ajutor managerilor în adoptarea unor decizii privind poluarea apei.

Agregarea la diferite niveluri ale ierarhiilor stabilite conduce la obținerea unor rezultate de sinteză prin operații de roll-up. De asemenea, operațiile de feliere și decupare de cubulețe (slice and dice) sunt utile ca răspunsuri la întrebări ce iau în considerare mai multe dimensiuni.

Calculul riscului prin măsurile definite dau o imagine asupra gradului de poluare a apei datorat unor poluanți specifici proceselor tehnologice din industria textilă.

Pentru a se constitui într-un instrument eficient pentru controlul poluării apei, modelul prezentat mai sus urmează a fi inclus într-un sistem suport de decizie [7], [8].

Mulțumiri

Autoarea lucrării mulțumește d-nei cercetător principal Floarea Pricop de la Centrul pentru Protecția Mediului LACECA SA., București, pentru materialele documentare și datele privind măsurătorile privitoare la poluarea apei în industria textilă.

Bibliografie

1. AIRINEI, D.: Depozite de date, Ed. Polirom, 2003.
2. FILIP, F.G.: Decizie asistată de calculator decizii, decidenți, metode și instrumente de bază, Editura Tehnică, București, 2002.
3. OLAP Council, OLAP: On – Line Analytical Processing (<http://www.dssresources.com/glossary/olaptrms.html>), 1997.
4. PENDSE, N.: What is OLAP? An Analysis of what the Increasingly Misused OLAP Term is Supposed to Mean. OLAP Report, Optima Publishing Ltd, (<http://www.olapreport.com/FASMI.HTM>), 2004.
5. PENDSE, N.: The Origins of Today's OLAP Products. The OLAP Report, Optima Publishing Ltd, (<http://www.olapreport.com/origins.htm>), 2005.
6. PENDSE, N.: What's in a Name? The OLAP Report, Optima Publishing Ltd, (<http://www.olapreport.com/name.htm>), 2002.
7. RĂDULESCU, C-Z.: Dezvoltare durabilă prin analiza și evaluarea riscului de poluare a apei în industria textilă, Raport de cercetare la proiect RELANSIN 2126, ICI, București, 2004.
8. RĂDULESCU, C-Z.: Datamart și OLAP în analiza și evaluarea poluării apelor industriale. În: Revista Română de Informatică și Automatică, vol. 15, nr. 1, ICI, București, 2005, pp. 11-18.