Particularitati ale securitatii bazelor de modele economice distribuite

Prof.dr. Ion IVAN, prep. Cristian TOMA, stud. Adrian VISOIU
Catedra de Informatica Economica, A.S.E Bucuresti

The data sets, the models and the procedures are stored in a distributed way in different locations. The models that present common characteristics are classified in models families and could be created with models generators. From security point of view there are special situations because of remote procedure call and security communications between different processes.

***Keywords**: economical models base, models generators, security of models base.*

Diversitatea modelelor economice

Economia moderna utilizeaza numeroase tehnici si metode de analiza, dintre care cele orientate spre latura cantitativa a evolutiei proceselor sunt cele mai importante. Construirea de modele economice are ca obiectiv analiza corelatiilor dintre factori de influenta si variabile, din care rezulta forme analitice ale dependentelor identificate. Exista o mare diversitate de modele economice, fiecare din ramurile stiintelor economice rezervându-le spatii de prezentare deosebit de largi.

Modelele directe sunt rezultatul perceptiei nemijlocite a variabilelor asociate factorilor de influenta. Daca se considera un proces P caruia i se asociaza variabila rezultativa y si care este determinat în evolutia sa de numerosi factori $F_1,F_2,\ldots,F_N$ carora li se asociaza variabilele $X_1,X_2,\ldots,X_N$, un model direct are forma: $y=\sum_{i=1}^{N} a_i X_i$, unde a_i este un coeficient de contributie, definit pe multimea {-1,1}. Astfel, evolutia stocurilor materialelor $M_1,M_2,\ldots,M_f$ se modeleaza prin constructia:
$y_j = a_1X_{1j}+a_2X_{2j}+a_3X_{3j}$, $j=1,2,\ldots,f$, unde:
y_j - stocul final al materialului j;
X_{1j} - stocul initial al materialului j;
X_{2j} - intrarile prin aprovizionare din materialul j;
X_{3j} - iesirile spre consum din materialului j.
Coeficientii de contributie au nivelurile $a_1=1$; $a_2=1$; $a_3=-1$.

Modele de optimizare liniara includ în structura lor functii de forma:

$\{\min/\max\} f(x)=\sum_{j=1}^{n} a_j c_j x_j$ unde:

n - numarul de produse;
a_{ij} - coeficientii de contributie definiti pe multimea {-1,0,1};
c_j - coeficientii de multiplicare;
x_j - nivelul variabilelor asociate factorilor, care influenteaza procesele de maximizare sau de minimizare.

Restrictiile modelelor liniare au forma:

$\sum_{j=1}^{m} a_{ij}a_{ij}x_j \le b_i, i=1,2,...,n$, unde:

a_{ij} - coeficientii de contributie pentru ecuatia i;
b_{ij} - nivel limitativ pentru definirea unui capat al intervalului de variabile pentru resursele utilizate;
a_{ij} - consumul unitar de resursa de tip i pentru a realiza o unitate de produs;
m - numarul de restrictii.

Modelele neliniare reprezinta o larga categorie utilizata în studierea interdependentelor dintre factori. Marea varietate de modele neliniare produce efecte variate în colectarea si dezvoltarea sistematizata a lor, întrucât în descrierea modelelor de acest tip trebuie definite reguli care sa conduca la descrieri corecte si la implementari cu grad de cuprindere deosebit de ridicat. Sunt situatii în care modelele neliniare, prin transformari convenabile sunt transformate în modele liniare.

De exemplu, modelul neliniar: $y=A\cdot X^{a}\cdot W^{ß}$, prin logaritmare conduce la structura $\log y = \log A + a\cdot\log X + ß\cdot\log Y$. Efectuând înlocuirile $y'=\log y$, $A'=\log A$, $X'=\log X$, $W'=\log W$, se obtine modelul liniar $y'= A'+ a\cdot X'+$

ß·W' care se poate rescrie în forma

$y' = \sum_{i=1}^{n} a_i b_i X_i$, unde $X_1=A'$, $X_2=A'$, $X_3=W'$, $b_1=1$, $b_2=a$, $b_3=ß$ si $a_1= a_2= a_3=1$.

Toate acestea conduc la ideea gasirii unor tehnici si metode de stocare si gestionare a diversitatii de modele economice, folosind software adecvat.

Seriile de date de intrare

Toate modelele economice utilizeaza date de intrare înregistrate conform unor proceduri stricte. În sistemul financiar-contabil exista documente tipizate, exista reguli de înregistrare si de calcul a indicatorilor – de exemplu, pentru completarea unei facturi exista o procedura care permite ca orice persoana care dezvolta un proces sa obtina un continut identic cu continutul documentelor completate independent de toate celelalte persoane care reproduc, de asemenea, identic procesul respectiv. Se construiesc serii de date care includ elemente de identificare si nivelurile variabilelor asociate factorilor.

Toate datele se sistematizeaza în tabele de forma X_Y, unde X – numarul de linii; Y – numarul de coloane.

Un bon de casa complet este de forma (n+3),4, întrucât contine:

- un numar de n linii corespunzatoare produselor achizitionate;
- o linie pentru înregistrarea valorii totale fara TVA;
- o linie pentru nivelul TVA total;
- o linie pentru suma de plata.

Cele trei coloane corespund – în varianta simplificata a unui bon de casa: coloana 1 – codului/denumirii produsului; coloana 2 – cantitatea achizitionata; coloana 3 – pretul produsului cumparat.

Seriile de date sunt input-ul unui model, si de regula acestea au un format recunoscut de model pentru a putea fi implementate prin produse software.

Generatoare de modele

Colectiile de modele economice evidentiaza existenta unor familii de modele. Se considera un model *M* care are în structura un set de variabile exogene. Trecerea la modelul *M'* se realizeaza prin includerea în setul de variabile a unui subset de variabile noi. Trecerea de la modelul *M*, prin eliminarea unei variabile, conduce la obtinerea unui model cu o structura mai simpla, '*M*. Atunci când modelului i se adauga variabile, procesul este de dezvoltare, iar operatorul este notat *D(.).D(M) = M'*. Când se trece de la un model *M* la un alt model '*M*, prin eliminarea unei variabile, procesul este de simplificare, iar operatorul este *S(.)*.

S(M)='M.

Se observa ca în cazul în care variabila adaugata este X_{i+1} modelul *M* contine deja variabilele $X_1, X_2, \ldots, X_i$.

$D(M)_{Xi+1} = M'$; $S(M)_{Xi+1} = 'M$

iar $S(D(M)_{Xi+1})_{Xi+1} = M$.

În cazul în care prin simplificare informatia obtinuta din model este în continuare relevanta, procesul se numeste rafinarea modelului.

Generatoarele de modele liniare reprezinta mecanisme deosebit de importante pentru obtinerea de modele economice reprezentative. Se considera variabilele independente X_1, $X_2, \ldots, X_n$ si variabila dependenta *Y*.

Practica economica a condus, în general, la elaborarea de modele liniare pentru ca:

- fenomenele studiate urmaresc o dependenta liniara;
- metodele de estimare a parametrilor sunt uzuale pentru aceste tipuri de modele;
- interpretarea rezultatelor este usurata daca sunt luate în calcul ipotezele de liniaritate.

Se construiesc modele cu o variabila, de forma: $y^{(k)} = a_1^{(k)} x_i + a_0^{(k)}$, $i=1,2,\ldots,n$ iar k este un numar de ordine al modelului.

Se construiesc modele cu doua variabile:

$y^{(k)} = a_i^{(k)} x_i + a_j^{(k)} x_j + a_0^{(k)}$, $i=1,2,\ldots,n$, i?j.

În acelasi fel se construiesc modele cu trei, patru variabile, iar cel mai complet dintre modele include cele n variabile independente Pe lânga generatoarele de modele liniare, se regasesc si alte generatoare pentru alte tipuri de modele.

Proceduri folosite pentru utilizarea de modele

Pâna la utilizarea unui model economic, este necesara parcurgerea a numerosi pasi.

Procedurile destinate estimarii coeficientilor

au ca date de intrare: lungimea seriilor de date asociate factorilor de influenta si variabilelor rezultative; seriile de date propriu-zise; structura modelului pentru care se estimeaza coeficientii, codul asociat metodei de estimare utilizate. Rezultatele oferite de aceste proceduri sunt:

- sirul de coeficienti estimati;
- nivelul sumei patratelor diferentelor dintre nivelurile reale si cele estimate ale variabilei rezultative;
- valorile asociate rezultatului testarii unor ipoteze statistice privind calitatea estimatorilor;
- informatiile privind derularea estimarii.

În cazul în care sunt utilizate metode de estimare care impun unele restrictii asupra seriilor de date de intrare ale modelului, înainte de derularea algoritmului de estimare se verifica ipotezele referitoare la acele restrictii. Daca sunt verificate ipotezele se lanseaza procesul de estimare. În caz contrar se trece la utilizarea altor algoritmi de estimare, liberi de aceste restrictii. Dupa utilizarea acestor proceduri, trebuie analizate informatiile returnate de proceduri pentru a valida modelul în care s-a derulat procesul de estimare si pentru a utiliza estimatorii într-un mod consistent, în raport cu ansamblul activitatilor de proiectare, analiza, realizare si utilizare curenta a modelelor economice.

Procedurile pentru operare pe seturi de date vizeaza realizarea:

- preluarea seturilor de date;
- omogenizarea seriilor de date prin interpolare, extrapolare si prin transformari elementare;
- concatenarea seturilor obtinând seturi de date cu un numar sporit de serii de date;
- extensia seturilor de date pentru a obtine serii cu un numar mai mare de termeni;
- asigurarea comprehensibilitatii datelor prin aplicarea de coeficienti de transformare termenilor;
- agregarea seturilor de date pentru obtinerea de indicatori noi.

Datele de intrare contin numarul seriilor de date, lungimile seriilor de date, tipologiile de prelucrari. Rezultatele obtinute se concretizeaza în noile serii de date, lungimile noilor serii de date, destinatiile de stocare a seriilor; informatii privind calitatea procesului de operare.

Procedurile pentru generarea de date de test au menirea de a pregati date de intrare care sa fie utilizate pentru studierea modelelor. Exista numeroase situatii în care structurile de date existente sunt incomplete, desi sunt cunoscute tendintele de evolutie a fenomenelor si limite de variatie a nivelurilor variabilelor asociate. De asemenea, lungimile seriilor de date existente la un moment dat sunt insuficiente pentru a produce estimari de calitate.

Pentru a pregati un model economic, trebuie generate seturi de date si studiate proprietatile modelului folosind seturile de date generate. Se simuleaza în acest fel comportamentul modelului. Se utilizeaza proceduri pentru generarea de numere pseudoaleatoare, care urmeaza diferite legi de distributie.

Particularitati ale securitatii bazelor de modele

Este evident ca modelele dezvoltate de diferite grupuri de cercetare pot fi stocate în baze de date sub forma de baza de modele. Un scenariu coerent aplicat într-un mediu distribuit este urmatorul: un proces client trimite datele de intrare unui proces ce implementeaza o anumita procedura; conform procedurii datele sunt agregate si transformate si apoi transferate unui proces ce implementeaza un model economic; procesul aplica acel model ales de procesul client iar rezultatele le paseaza procesului ce implementeaza procedura care conform pasilor din procedura va transfera rezultatele la procesul client (figura 1).

În ipoteza în care seturile de date, modelele, generatoarele de modele si procedurile se afla pe dispozitive diferite – procesul client poate rula inclusiv pe un telefon mobil sau PDA, apar probleme de securitate. De exemplu un client ce are un anumit set de date, doreste sa primeasca rezultatele aplicarii unui anumit model economic pe datele de intrare. Dar este evident ca nu orice client poate folosi un anumit model, existând restrictii de autentificare, integritate, confidentialitate si non-repudiere asupra serviciului oferit de procesele ce implementeaza procedura si executia modelului.

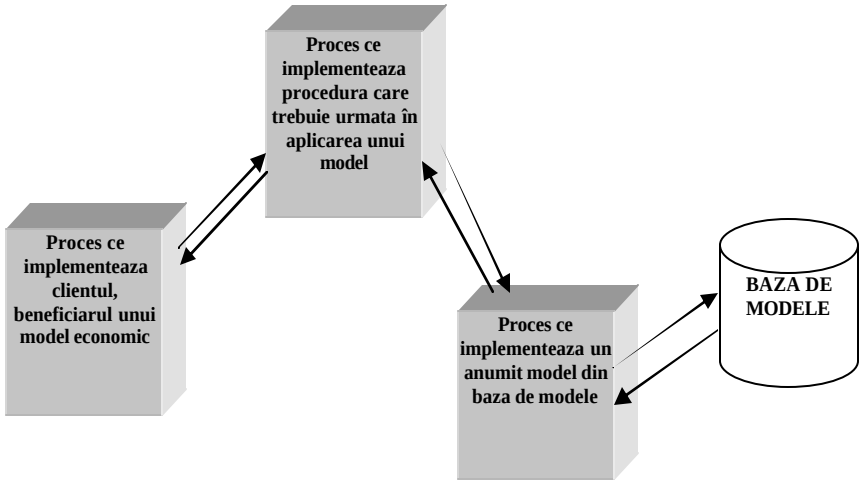

Fig. 1. Procese distribuite în retea pentru aplicarea unui model

Protocolul de comunicare între diferite componente nu este detaliat dar se fac anumite observatii. În general protocoalele de comunicatie dintre diferitele procese din retea sunt de tip text XML – eXtended Markup Language. Comunicarea este securizata conform arhitecturii prezentate în [IVAN02] unde pentru semnatura digitala autentificarea se foloseste ca functie hash MD5, iar ca algoritm cu chei publice RSA. Între diferitele procese se folosesc protocoale text XML deoarece procesele sunt rezidente în retele eterogene care prezinta particularitati diferite cum ar fi: diferite proxy-uri, gateway-uri si firewall-uri, diferite arhitecturi si stive de protocoale – TCP/IP, Token Ring, WAP. Pentru apelul metodelor la distanta, RPC – Remote Procedure Call sau RMI – Remote Method Invocation, se foloseste protocolul SOAP – Simple Object Access Protocol în care parametrii catre proceduri se transmit XML peste HTTPS - Hyper Text Transfer Protocol Secure, iar rezultatul este întors tot XML peste HTTPS.

Pentru a include facilitati de securizare – oricare doua procese implicate în comunicatie au certificate digitale X.509 v3 corespunzatoare – se modifica în modul urmator: clientul apeleaza o metoda la distanta a procesului ce implementeaza procedura prin SOAP peste HTTPS; procesul ce implementeaza procedura primeste parametrii, adica seriile de date de la client; procedura clientului transforma seriile de date conform unor reguli si le trimite în mod securizat la procesul ce implementeaza un anumit model – de data aceasta nu a mai fost apel de procedura la distanta; procesul alege un model din baza de modele si-l aplica datelor primate; rezultatul fiind returnat pe calea inversa.

În proiectarea protocoalelor securizate de comunicatie între procese trebuie sa se tina seama de tipul de distribuire a modelelor de date, procedurilor si seturilor de date, considerându-se urmatoarele posibilitati:

- estimarea coeficientilor corespunzatori modelelor folosind ecuatiile clientului – trimise ca date procesului ce implementeaza procedura - si cu seturile de date tot ale clientului;
- estimarea coeficientilor corespunzatori modelelor din bazele existente construite de diferiti producatori, dar folosind seturile de date ale clientului;
- construirea de modele cu ecuatii simultane în baze de modele diferite;
- efectuarea de prelucrari pe seriile de date pentru a obtine serii omogene echidistante si de lungime impusa prin extrapolare, iar acolo unde lipsesc termeni interiori se procedeaza la interpolare – aceasta caracteristica nu este neaparat necesar sa fie implementata de procesul procedura.

Protocoalele de comunicare momentan sunt bazate pe comunicare de mesaje sincrone si securizate la nivelul aplicatiei din stiva de protocoale ISO/OSI. Toate aceste consideratii sunt facute în ideea de a nu restrictiona

procesele implicate la un anumit tip de retea sau de dispozitive.

Concluzii

Momentan nu exista un standard care sa reglementeze arhitectura pentru folosirea unui model economic asupra unor serii de date. De asemenea nu exista standard care sa specifice ce tip de protocoale de comunicatie se pot folosi si modul de transfer a datelor între procese ce au functii diferite. Cu toate acestea în arhitectura din figura 1 se recomanda folosirea unor standarde recunoscute pentru protocoalele de comunicatie – SOAP peste HTTPS pentru apelul procedurilor la distanta; semnatura digitala XML pentru transmiterea datelor împreuna cu algoritmul Rijndael pentru asigurarea confidentialitatii; CORBA. Arhitecturile pentru comunicarea interprocese pentru executarea de calcul rapid la procesele ce implementeaza modele.

Bibliografie

[IVAN02] Ion Ivan, Paul Pocatilu, Marius Popa, Cristian Toma, „*Semnatura electronica si securitatea datelor în comertul electronic*", Informatica Economica Nr. 3/2002, Bucuresti 2002.

[Lyu96] M. R. Lyu (Ed.), *Handbook of Software Reliability Engineering*, IEEE Computer Society Press, 1996.

[Mathur97] S. Krishnamurthy, A.P. Mathur, „*On the estimation of reliability of a software system using reliability of its components*", Proceedings of the 8th IEEE International Symposium on Software Reliability Engineering (ISSRE'97), pp.146 – 155, Nov. 1997.