

# Analiza chat-urilor CSCL folosind pachetul R

**Iulia Badea**

Universitatea “Politehnica” București

313 Splaiul Independenței, București,  
060032, Romania

juliebadea@gmail.com

**Stefan Trausan-Matu**

Universitatea “Politehnica” București

313 Splaiul Independenței, București,  
060032, Romania

Institutul de Cercetări în Inteligență artificială

Calea 13 Septembrie nr. 13, București,  
România

stefan.trausan@cs.pub.ro

## REZUMAT

Lucrarea prezintă o abordare de analiză a textelor ce utilizează seriile de timp concomitent cu mineritul textelor (*text mining*), adică extragerea de cunoștințe din texte, cu scopul de a analiza tipuri de texte care sunt strâns legate de factorul timp, al căror conținut poate fi reprezentat pe o axă temporală, cum ar fi log-urile chat-urilor. Aceste tehnici aplicate pe documente de tip chat permit determinarea de corelații între cuvintele care sunt întâlnite cel mai frecvent în cadrul dialogului, și care, datorită frecvenței mari de apariție reprezintă teme de discuție în chat, precum și determinarea de corelații între intervențiile participanților la dialog. Acestea pot fi reprezentate folosind modelul seriilor de timp, în care coordonata timp este reprezentată de momentul scrierii replicii, iar valoarea în punctul respectiv o reprezintă replica în sine. Pornind de la această reprezentare, pot fi cu ușurință calculate corelații între ritmicitățile intervențiilor cu grad ridicat de apariție. Cercetările folosesc pachetul R, care oferă facilități pentru prelucrările statistice.

## Cuvinte cheie

Serii de timp, CSCL (învățare colaborativă asistată de calculator), polifonie, ritmul conversației, interacțiuni în chat.

## Clasificare ACM

I.2.7. ARTIFICIAL INTELLIGENCE: Natural Language Processing.

## INTRODUCERE

Conversațiile online (chat-urile) sunt unele dintre cele mai folosite și eficiente mijloace de comunicare din punctul de vedere al învățării colaborative asistată de calculator (CSCL – Computer Supported Collaborative Learning). Au fost deja dezvoltate mai multe aplicații pentru analiza interacțiunii în chat-uri. Ceea ce aduce nou abordarea de la care pleacă lucrarea de față este modelul teoretic impus de filozoful și lingvistul rus Bakhtin, cunoscut drept modelul polifonic [18], care a fost aplicat pentru CSCL [19-21-22].

Acest model se bazează pe o analogie cu muzica, unde timpul, în special prin componenta sa, ritmul, este principala dimensiune în funcție de care se realizează analiza textului [21-22].

Analiza seriilor de timp este o unealtă cu un real potențial în învățarea asistată de calculator pentru că astfel se pot

genera statistici bazate pe dinamica textului, cum este succesiunea replicilor pentru chat-uri. Deși seriile de timp sunt o unealtă rareori folosită pentru a face asocieri și comparații între sisteme cum sunt chat-urile, care pot evolua ușor și imprevizibil în timp, acestea pot reprezenta o posibilitate de a analiza fenomene și interacțiuni pentru care marca temporală are o deosebită importanță.

Comunicarea prin chat-urile online pentru învățarea colaborativă asistată de calculator reprezintă un tip de interacțiune în care, pe lângă dimensiunea temporală esențială, ritmul conversației și polifonia discursului, ambele inspirate din sfera muzicală, sunt extrem de importante [19-22].

Lucrarea propune o analiză a unei colecții de log-uri de chat. Acest “corpus” de fișiere este obținut ca rezultat al activităților colaborative realizate de studenții Departamentului de Calculatoare din cadrul Universității Politehnica București în cadrul unor cursuri [23].

Articolul cuprinde o secțiune ce prezintă, în linii mari, aparatul teoretic care a stat la baza realizării analizei propuse, urmată de o parte în care este descris modelul polifonic Bahtinian ce a determinat această abordare.

Următoarea parte al lucrării se întinde pe parcursul mai multor secțiuni și acoperă informații atât despre preprocesarea colecției de texte studiate, tehnicile împrumutate din prelucrarea limbajului natural (NLP) aplicate pe texte, cât și despre felul în care este evaluată implicarea fiecărui participant la chat. Lucrarea se încheie cu o secțiune de concluzii și referințe.

## APARAT TEORETIC

Reintroducerea anumitor elemente semantice la intervale prestabilite de timp, precum și repetarea unor termeni în mod periodic constituie două tehnici prin care se creează ritm și se induce ideea de muzicalitate în texte [20].

Folosirea tehnicilor de prelucrarea limbajului natural pentru analiza activităților de tip colaborativ este un domeniu important în Inteligența Artificială și este utilizat în învățarea colaborativă asistată de calculator sub denumirea de *Learning Analytics*, deoarece ea reprezintă o metoda de a primi feedback pentru îmbunătățirea și eficientizarea activităților de tip colaborativ. Dimensiunea temporală reprezintă un aspect important al colaborării, deoarece reflectă dinamica dialogului [16].

În unele cazuri, repetițiile reprezintă indicatori ai similarităților sau, în mod contrar, ai diferențelor care

conduc la interrelaționări între conceptele care apar în texte. În acest context, aparițiile frecvente ale unor termeni pot contribui la ridicarea nivelului literar al unui text, prin adăugarea de noi înțelesuri și stabilirea de relații noi între entități. Acest tip de date poate fi studiat utilizând metode statistice, una dintre acestea fiind aleasă și în această lucrare: analiza “evenimentelor” discrete din chat-uri în funcție de momentul în care au loc, înscriindu-se pe o axă temporală și generând, astfel, serii de timp [16].

#### MODELUL DIALOGISMULUI LUI BAKHTIN

Dialogismul lui Bakhtin poate fi considerat un model teoretic pentru CSCL [19-22]. Acesta susține că în orice text, fiecare cuvânt reprezintă ecourile mai multor replici anterioare, de aici și ventrilogismul sau multivocalitatea întâlnite la Bakhtin. Replicile pot genera voci care să producă la rândul lor ecouri. Trăușan-Matu afirmă că, între muzica polifonică și analiza conversațiilor dintr-un chat, există un grad de similaritate [19-21]. Pornind de la această idee se poate afirma că fiecare voce reprezintă un fir al discuției într-un text dialogat. La Bakhtin nu există separația clară între emițător și receptor, în locul unui transfer, comunicarea bakhtiniană este un proces multivocal, în care vocile interlocutorilor și nu numai ale lor, se interanimă, se întrepătrund, compenetrează (cum, în plus față de Bakhtin spunea și Noica în 1986 [24]).

Mai mult decât atât, vocile pot să interacționeze între ele creând o adevărată polifonie, adică o multitudine de voci ale unor conștiințe independente, necontopite, dar care se interanimă [19-21]. Astfel, pentru analiza discursului într-un chat se verifică dacă vocile sunt legate între ele și modul în care ele interacționează din punct de vedere polifonic.

Dialogistica stă la baza, printre altele, și a paradigmei învățământului colaborativ sprijinit de calculator. Realizările activităților de tip colaborativ bazate pe chat (dialog online) și, în particular, activitățile de învățare prin colaborare depind de gradul de inter-animare polifonică (intermixare sau interpenetrare, în alte traduceri așa cum le numea M. Bakhtin) a mai multor voci în discursuri [19]. Începând cu ideile lui Mikhail Bakhtin despre multivocalitatea unui text sau a unei comunicări și despre polifonie, putem spune chiar că, și dacă învățăm individual, participăm la o polifonie vocală, la un discurs dialogat, intrăm într-o relație de colaborare cu, de exemplu, autorul cărților pe care le studiem. Din perspectiva opusă, chiar dacă autorul nu va auzi probabil niciodată vocea noastră actuală, simplul fapt că autorul scrie pentru cititori reprezintă un potențial dialog în care acesta intră în contact cu vocile lor. Chiar dacă învățarea este individuală, colaborarea rămâne o caracteristică foarte importantă, împărtășirea ideilor cu tot ce implică acest lucru, de exemplu, acordul sau dezacordul urmate de negocieri. Oricum, dacă participăm la o învățare de tip colaborativ unde auzim vocile actuale, nu cele întruhipate în cărți, procesul devine mai direct, și intrăm în fenomenul polifonic.

Tehnologia informației și a comunicării oferă acum noi posibilități pentru colaborare, un bun exemplu fiind faptul că în chat-uri putem să folosim mult mai ușor un discurs

cu mai multe fire de execuție, similar contrapunctului în muzica clasică, decât în conversațiile clasice.

De asemenea, poate fi generat și un feedback al conversației, cum este cazul lucrării de față, cu statistici care să includă: lista celor mai importante (utilizate, discutate) concepte într-un chat / forum, acoperirea celor mai importante concepte specificate de cel care a inițiat conversația, cele mai importante intervenții ale fiecărui participant, un scor pentru fiecare participant la conversație.

#### ACHIZITIA DE DATE SI PROCESAREA LOR

Ideea utilizării seriilor de timp este necesară pentru a studia corelațiile între aparițiile unor concepte în secvențele de texte studiate și implicarea la dialog a participanților la chat în momentul introducerii de noi concepte, la anumite intervale de timp [15]. Analiza propusă urmează o serie de pași începând cu realizarea de log-uri pentru chat-urile studiate și continuând cu preprocesarea lor, descrisă într-una dintre secțiunile ce vor urma. Log-urile de chat utilizate au fost cele realizate în timpul activităților de învățare de tip colaborativ desfășurate în cadrul Universității Politehnica București.

Pentru a detecta pattern-uri, șabloane sau repetiții prin care este indusă ideea de ritm, primul pas constă în preprocesarea textelor observate folosind tehnici din sfera prelucrării limbajului natural: extragerea de informații, recunoașterea subiectelor sau a temelor principale, recunoașterea principalelor entități, gruparea conceptelor din texte, precum și identificarea principalilor participanți la dialog și generarea de serii de timp în funcție de intervențiile la chat ale fiecăruia.[14]

#### Parsarea XML-ului

Formatul fișierului de intrare este de tip XML. Într-o astfel de reprezentare participanții sunt ușor de identificat. Numele lor va apărea la începutul fișierului, sub următoarea formă:

```
<Participants>
<Person nickname="Mona-chat"/>
<Person nickname="stefan-blog"/>
<Person nickname="corina-forum"/>
<Person nickname="diana-wiki"/>
<Person nickname="cristi-wave"/>
</Participants>
```

O replică va fi memorată în fișier astfel:

```
<Turn nickname="Mona-chat">
<Utterance genid="29" ref="27"
time="11:50:58">Let's think od some
activities... and then decide which
technology is better</Utterance>
</Turn>
```

Fiecare din date reprezintă una din următoarele:

- valoarea *nickname*, numele participantului căruia îi aparține replica;
- valoarea *genid*, id-ul asociat replicii;
- valoarea *ref*, id-ul replicii la care se face referința, valoarea este 0 dacă nu se refera nici o replica;

- valoarea *time*, momentul de timp la care este scrisa replica;
- valoarea *dintre taguri*, replica în sine.

Prima etapă în preprocesarea codului din fișierul XML reprezentat de fiecare log de chat, constă în identificarea principalelor tag-uri care conțin: *nickname*-ul participantului la chat, timpul intervenției, precum și replica. Toate aceste lucruri au fost realizate cu ajutorul funcțiilor din pachetul *XML*, respectiv *textmining*, din limbajul *R*.

Pentru prelucrarea datelor discuției, este necesară extragerea replicilor din fișierul XML în care s-a făcut exportul. Pentru implementarea parserului a fost folosit DOM (Document Object Model). DOM reprezintă o convenție de reprezentare și interacțiune cu obiectele din HTML, XHTML și XML. Este independent de platformă și de limbajul de programare. Folosind această structură este suficient să se apeleze funcțiile de tip parse pentru a se face parsarea fișierului XML și interogarea obiectelor dintr-un DOM. Prima etapă constă în determinarea participanților la chat.

Participanții apar în taguri de forma: `<Person nickname="Mona-chat"/>`. Se caută în DOM-ul nou format elementele având tagul *Person* și va fi întorsă o listă a nodurilor care conțin numele participanților. Pentru fiecare nod din acea listă, se determină valoarea atributului *nickname*. Din valorile tuturor atributelor se va forma o listă cu numele tuturor participanților.

```

> source("C:\\Users\\julie\\Documents\\chat-nou.R")
[1] "nr participanti"
[1] 5
[1] "participanti"
[[1]]
[1] "Mona-chat"

[[2]]
[1] "stefan-blog"

[[3]]
[1] "corina-forum"

[[4]]
[1] "diana-wiki"

[[5]]
[1] "cristi-wave"

> |
    
```

Figura 1. Participanții la chat

După ce s-au identificat participanții la discuție urmează determinarea tuturor replicilor pentru fiecare participant în parte. Acestea vor fi memorate sub forma unei liste care va conține la rândul său alte liste care vor conține replicile participanților. Pentru identificarea replicilor se determină elementele având tagul *Turn*. Se formează o listă de noduri corespunzătoare. Pentru fiecare astfel de nod, se găsește primul copil al său dintre copiii determinați și se rețin valorile atributelor: *nickname*, *genid*, *ref si time*, precum și valoarea nodului copil de tip text, care reprezintă replica. Se elimină replicile de tipul “joins the room” sau

“left the room”, deoarece nu au nici o relevanță în evaluarea sesiunii de chat.

```

$'cristi-wave'
[1] "joins the room"
[2] "hello"
[3] "We are supposed to be in company , right ?"
[4] ""
[5] "right"
[6] "yes...we should think of some activities ... like a board meeting where some of us can't be there ...what tech should i use ?"
[7] "wave has a presentation template where videochat is included ..with multiple features"
[8] "not if our customer require a fast response"
[9] "% videochat with possibility to draw in real time some charts for example ... here wave is very helpful"
[10] "wave...videochat requires some bandwidth"
[11] "nowadays a fast internet connection is everywhere :)"
[12] "undoubtedly"
[13] "true .... wiki = how to"
[14] "not if the project is an open source :)"
[15] "none"
[16] "it was the end for my sentence above :)"
[17] "for small projects ,yes"
[18] "not for a large one ... use cases , lots of pdf files ... probably wiki"
[19] "yes...sometime the client wants to be in touch with all the stages"
[20] "I don't think is going to happen :)"
[21] "yes...chat is sometimes unreliable"
[22] "meeting board"
[23] "agree"
[24] "I see a lot of blogs for selling stuff ... I agree"
[25] "and posting some images for example"
[26] "for selling the best solution is something like ebay , but this is not our topic now"
[27] "good point"
[28] "at least a voice chat ... I have not heard so far about selling something , just chatting :)"
[29] "I think so"
[30] "yes ... in a forum topics are very well delimited .... it is very easy to read it"
[31] "the only wrong question is the one not asked :)"
[32] "I think a chat room is suitable in this case"
[33] "is about human resources i think"
[34] "yes"
[35] "unless we have a bot with some AI :)"
[36] "true..but none of the developers would want to answer some questions in a chat ... their job is to develop , not client support"
[37] "full time support means more money"
[38] "more money for the company ... this kind of support is more expensive"
    
```

Figura 2. Replicile unui participant la chat

După ce au fost determinate listele de replici pentru fiecare participant în parte, se trece la partea de procesare a fiecărei replici în parte folosind tehnici de procesare a limbajului natural. Pentru determinarea listelor de cuvinte folosite de către fiecare participant în parte, s-au folosit tehnici NLP (natural language processing): extragerea rădăcinii cuvintelor (stemming) și detectarea părților de vorbire.

### CONTRIBUTIA UNUI PARTICIPANT LA O SESIUNE DE CHAT

În cadrul unei sesiuni de învățare prin colaborare este importantă analiza contribuției fiecărui participant la discuție. Această analiză ar putea porni de la cât de mult a influențat studentul conversația. În funcție de implicarea participantului în discuție se poate determina cât de puternic a influențat acesta sesiunea de chat. O astfel de analiză ar putea porni de la determinarea numărului de cuvinte dintre cele specifice temelor de discuție alocate pentru dezbateri, folosite de participant în cadrul discuției. Dacă există un număr mare de replici care să conțină astfel de cuvinte, atunci se poate continua prin obținerea de indicii asupra faptului că participantul a avut replici în care si-a exprimat acordul, dezacordul sau a pus întrebări. Toate acestea ne indică cât de puternic a fost influențată sesiunea de învățare de către un participant.

### METODE DE EVALUARE

În cadrul proiectului s-a luat în considerare că evaluarea contribuției fiecărui participant în cadrul unei sesiuni de învățare se face pe baza numărului de cuvinte folosite din cadrul subiectelor alocate pentru sesiunea de învățare și pe baza unui algoritm de evaluare a colaborării. Ideea de la care se pleacă este că fiecare student are asociată o

poziție/temă de discuție, pe care o cunoaște înaintea începerii sesiunii de învățare. Fiecare temă va reprezenta un fir de discuție în cadrul sesiunii de chat. Discuția va avea loc într-un mediu în care se pot realiza legături explicite între replici, acest lucru făcându-se prin crearea de referințe între replici. În Figura 3, este dată ca exemplu o mostră dintr-o sesiune de chat.

```
<?xml version = "1.0" ?>
<Dialog team="34">
<Participants>
<Person nickname="Mona-chat"/><Person nickname="stefan-blog"/><Person nickname="corina-forum"/><Person nick
" cristi-wave"/></Participants> <Topics>
<Body>
<Turn nickname="Mona-chat">
<Utterance genid="1" ref="0" time="11:43:05">joins the room</Utterance>
</Turn>
<Turn nickname="corina-forum">
<Utterance genid="2" ref="0" time="11:45:01">joins the room</Utterance>
</Turn>
<Turn nickname="corina-forum">
<Utterance genid="3" ref="0" time="11:45:06">Hello</Utterance>
</Turn>
<Turn nickname="Mona-chat">
<Utterance genid="4" ref="0" time="11:45:08">Hello, Corina!</Utterance>
</Turn>
<Turn nickname="diana-wiki">
<Utterance genid="5" ref="0" time="11:45:18">joins the room</Utterance>
</Turn>
<Turn nickname="cristi-wave">
<Utterance genid="6" ref="0" time="11:45:23">joins the room</Utterance>
</Turn>
<Turn nickname="Mona-chat">
<Utterance genid="7" ref="0" time="11:45:28">Hello, Diana!</Utterance>
</Turn>
<Turn nickname="Mona-chat">
```

Figura3. Mostră dintr-un fișier de chat

## REZULTATE

Până acum, cercetarea s-a concentrat pe analiza unui fișier de tip log pentru a clasifica și pentru a analiza activitățile de învățare de tip colaborativ, în timp ce în această lucrare accentul este pus pe dimensiunea temporală și pe înscrierea pe o axa cronologică a evenimentelor din chat și pe selectarea celor mai importanți termeni în conversație.

De asemenea, alte informații au fost obținute prin simpla parsare a fișierelor-log de chat și aplicarea funcțiilor specifice pentru a procesa arborele, incluse în pachetele din limbajul R: numele participanților la chat, numărul total de participanți, numărul total de replici, replicile grupate pe fiecare participant, cel mai activ participant din punctul de vedere al numărului de replici.

În final, timpul pentru fiecare replică a unui participant a fost obținut și introdus într-o serie de timp. Seriile de timp au avantajul că poate fi calculat timpul mediu între două replici consecutive ale aceluiași participant, astfel evidențiindu-se cel mai activ participant la chat sau alternanța rolurilor participanților la chat.

Fiecare student are atribuită o temă de discuție. Participantul la discuție poate spune o replică care se referă la unul sau mai multe teme alocate altor participanți la discuție (cv), sau poate face referire la o altă replică care poate conține teme de discuție (rv) [25]. Pot apărea diferite situații când sunt găsite replici de tipul (cv) sau participantul se referă la alte replici (rv). Astfel participantul:

- vorbește despre propria temă;

- face referire la o replică care conține tema alocată lui;
- vorbește despre aceeași temă ca și cea din replica referită;
- vorbește despre o altă temă decât propria temă;
- face referire la o replică care nu conține o tema alocată lui;
- vorbește despre altă temă decât cea care este abordată în replica la care face referire.

Realizând o analiză al numărului de apariții al situațiilor descrise mai sus și a rapoartelor acestora se poate obține o imagine asupra contribuției fiecărui participant, atât din punct de vedere colaborativ cât și din punctul de vedere al inter-anietației. De exemplu, o contribuție bună ar trebui să fie ilustrată de rapoarte relativ egale între replicile orientate către propriul subiect și cele care se referă la alte subiecte. Un astfel de raport poate fi  $nriv \neq cv/nriv = cv$  care în cazul în care se apropie de 1 ar indica faptul că participantul a avut o colaborare bună în cadrul sesiunii de chat [25].

## IMPLEMENTARE

Aplicația are rolul de a analiza o sesiune de colaborare prin învățare, sesiunea considerată fiind de tip chat. Analiza constă în identificarea firelor de discuție în cadrul sesiunii de chat precum și identificarea gradului de implicare al fiecărui participant din punct de vedere colaborativ. Aplicația este dezvoltată pentru a evalua orice tip de chat care urmează următorul șablon: exista un număr de participanți, fiecare având asociată o temă de discuție, toate temele fiind din același domeniu. Un participant va susține un anumit subiect, aducând argumente pro acestuia și argumente împotriva celorlalte subiecte. Nu este necesar ca toți studenții implicați să aibă asociate teme diferite. După rularea aplicației, se vor observa cuvintele din chat repetate cel mai des în timpul discuției și numărul lor de apariție, acestea vor fi grupate după tipul de părți de vorbire substantive, verb, adverb sau adjectiv. Pe baza acestor liste se vor putea identifica firele de discuție apărute în cadrul sesiunii de învățare. În plus prin aplicarea algoritmului de detectare a colaborării se poate determina gradul de colaborare al fiecărui participant.

## CONCLUZII

În cadrul acestei lucrări au fost prezentate diferite teorii care au ca scop evaluarea participanților la o sesiune de învățare prin colaborare. O nouă teorie, inspirată din ideile lui Bakhtin, a fost propusă pentru a explica și pentru a evalua colaborarea și interacțiunea în dialoguri [19,20]. Ideea ei principală este considerarea îmbinării thread-urilor (firelor) în discursuri asemenea tehnicii contrapunctului în muzica clasică. Vizualizarea grafică și diversele măsurători sunt calculate folosind o varietate mare de tehnici din analiza limbajului natural pentru analiza nivelurilor lexical, respectiv semantic al discursului. Primele experimente cu sistemul implementat au demonstrat că modelul polifonic ușurează dezvoltarea anumitor algoritmi și implementarea unui sistem oferă feedback participanților la dialog [22]. Cercetarea propusă urmărește două aspecte: primul teoretic, bazat pe studiul dialogisticii polifonice, iar al doilea practic, prin

implementarea unui sistem care analizează textele din sfera dialogismului, dezvoltând tehnicile actuale de analiză din procesarea limbajului natural și folosind unele tehnici deja existente.

Plecând de la teoria dialogismului lui Bakhtin, împreună cu folosirea tehnicilor de procesare a limbajului natural a fost conturat un model teoretic capabil să determine firele de discuție generate în cadrul sesiunii de învățare. Scopul proiectului a fost de a dezvolta o aplicație capabilă să detecteze firele de discuție generate în cadrul unei sesiuni de colaborare prin învățare de tip chat pentru discuții în limba engleză. Fiecărui participant la discuție îi este atribuit un subiect de discuție, aceeași temă putând fi susținută de mai mulți participanți.

Pe baza teoriei dialogismului lui Bakhtin susținătorii aceluiași subiect formează o voce. Aceste voci intră într-un proces de inter-animare generând astfel firele de discuție. Pentru detectarea firelor de discuție s-au folosit tehnici de procesare a limbajului natural precum *part of speech tagging* și *stemming*, care au contribuit la determinarea numărului de apariții al fiecărui cuvânt în parte. Plecând de la numărul de apariții al temelor abordate în cadrul sesiunii de învățare se pot stabili legături implicite între replici generând astfel firele de discuție. Pe lângă detectarea firelor de discuție aplicația implementează un algoritm de determinare a nivelului de implicare din punct de vedere colaborativ al fiecărui participant. Acest lucru poate fi determinat realizând o analiză asupra numărului de replici folosite în cadrul sesiunii de învățare. Plecând de la aceste replici se analizează numărul de replici care se referă la subiectul alocat participantului respectiv și la numărul de replici care se referă la subiectele alocate celorlalți participanți.

Realizând un raport al replicilor care s-au referit la subiectul alocat asupra replicilor care s-au referit la subiectele alocate celorlalți participanți s-a putut determina nivelul de colaborare al fiecărui participant în parte [25]. Aplicația implementată pe baza acestei idei oferă posibilitatea analizării fișierelor de tip .xml care respectă un anumit format.

#### REFERINTE

1. Baek E & Brock W , 1992. A general test for nonlinear granger causality bivariate model. Working paper, Iowa state university and University of Wisconsin, Madison
2. Fujihara R A and M Mougoue 1997. An examination of linear and nonlinear causal relationship between price volatility and volume in petroleum futures markets. Journal of Futures Markets 17, p385-416.
3. Hadsell L. 2006. A TARARCH examination of the return volatility – volume relationship in electricity futures. Applied Financial Economics, 16, 893 – 901.
4. Labys, W.C. 2006. Modeling and Forecasting Primary Commodity Prices. Ashgate Hampshire, England.
5. Puri, T. & G C Phillippatos, 2008. Asymmetric volume – return relation and concentrated trading in LIFFE futures, European Financial Management, vol 14, no 3, 528 – 563.
6. Kim, D. I., Sudderth, E. B. (2011), “The doubly correlated nonparametric topic model. Neural Information Processing Systems”.
7. I-A Chounta, N Avouris – “Time Series Analysis of Collaborative Activities”, HCI Group, University of Patras, Greece
8. R. Dahlhaus, M. Eichle - “Causality and graphical models in time series Analysis”, Universitat Heidelberg
9. T.Schümmer, J-W Strijbos, T. Berkel - “A New Direction for Log File Analysis in CSCL: Experiences with a Spatio-temporal Metric”
10. M. Kalisch, M. Martin, D. Colombo, M. H. Maathuis, B. Peter - “Causal Inference using Graphical Models with the R Package pcalg” - ETH Zurich
11. Y. Liu - Causality Analysis in Large-scale Time Series Data - Computer Science Department - Viterbi School of Engineering - University of Southern California
12. <http://hosho.ees.hokudai.ac.jp/~kubo/Rdoc/library/lmtest/html/grangertest.html>
13. Y. Liu, M. T. Bahadori – “A Survey on Granger Causality: A Computational View” - University of Southern California, October 22, 2012
14. Identifying Semantic Patterns: Time Series and Fourier Analyses Nancy M. IDE
15. Manu Kapur - Temporality matters: Advancing a method for analyzing 4 problem-solving processes in a computer-supported 5 collaborative environment 6
16. Irene-Angelica Chounta, Nikolaos Avouris: It’s all about time: towards a real-time evaluation schema for collaborative activities - HCI Group, University of Patras, Greece
17. Jeff\_B\_Cromwell\_MathWorks\_Presentation\_2012
18. Clark, K., Holquist, J.M., Mikhail Bakhtin, Harvard University Press, 1984
19. Trausan-Matu, S. (2010). The polyphonic model of hybrid and collaborative learning. In F. L. Wang, J. Fong, & R. C. Kwan (Eds.), Handbook of research on hybrid learning models: Advanced tools, technologies, and applications (pp. 466–486). Hershey: Information Science Publishing.
20. Trausan-Matu, S. (2012). Repetition as artifact generation in polyphonic CSCL chats. In Third Int. Conf. On Emerging Intelligent Data and Web Technologies (pp. 194–198). IEEE.
21. Trausan-Matu, S., Stahl, G., & Zemel, A. (2005). Polyphonic inter-animation in collaborative problem solving chats. Philadelphia: Drexel University, [http://mathforum.org/wikis/uploads/Stefan\\_Interanimation.doc](http://mathforum.org/wikis/uploads/Stefan_Interanimation.doc).
22. Trăușan-Matu, S, Rebedea, T, A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants, In A. F. Gelbukh (Ed.), 11th Int. Conf. Computational Linguistics and Intelligent Text Processing (CICLing 2010) (pp. 354–363). Iasi: Springer, 2010.
23. Stefan Trausan-Matu, Mihai Dascalu, Traian Rebedea, Alexandru Gartner, Corpus de conversatii multi-participant si editor pentru adnotarea lui, Revista Romană de Interactiune Om-Calculator, Vol.3, No.1, 2010, pp. 53-64, MatrixRom

Popovici, D.M., Iordache D.D. (Eds.)

24. Stefan Trausan-Matu, Dialogistica lui Bahtin: polifonie, ventrilogism, carnavalesc, inter-animare, în Constandache, G.G. (Ed.) Comunicarea - sugestie si influenta, Editura Sigma, pp. 39-44, (2009)
  25. Trausan-Matu S Experiencing, Conducting, Designing and Evaluating Polyphony in CSCL Chats.
- In: Spada H, Stahl G, Miyake N, Law N (eds) 9th Int. Conf. on Computer-Supported Collaborative Learning (CSCL 2011), Hong Kong, China, 2011. ISLS, pp 906–907