

Memory optimization strategies for linear mappings and indexation-based shared documents

M. Mumtaz Ahmad

Loria, INRIA Nancy - Grand Est, UHP Nancy - 1, IAEM Lorraine

November 14, 2011



Nancy-Université



General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

General introduction

- Speeding up computations
- Optimizing indexed communications

2 Sequential break down of operations

- Problem description
- Contributions

3 Decentralized collaborative editing system

- Problem description
- Contributions

Conclusions

- Summary
- Research directions



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

General introduction

- Speeding up computations
- Optimizing indexed communications
- Sequential break down of operations
 - Problem description
 - Contributions

Decentralized collaborative editing system

- Problem description
- Contributions

Conclusions

- Summary
- Research directions



Speeding up

computations

Speeding up computations

Higher performnace

 Modern computer technology has increased the demand for higher performance in all areas of computing.





Speeding up

computations

Speeding up computations

Higher performnace

• Modern computer technology has increased the demand for higher performance in all areas of computing.

Significant growth in

- Hardware performance
- Architecture evolution

Stress is raised on

- Compiler technology
- ry h directions

- ology
- Research communities





Speeding up computations

General

Speeding up computations

Optimizing indexed communications

Sequential break down of operations Problem description

Contributions

Decentralized collaborative editing system

Contributions

Conclusions Summary Research directions

Higher performnace

• Modern computer technology has increased the demand for higher performance in all areas of computing.

Significant growth in

- Hardware performance
- Architecture evolution

Stress is raised on

- Compiler technology
- 2 Research communities





.

General introduction

Speeding up computations

Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Micro-processor

- General-purpose PCs
- Servers to desktop, laptop
- Cell-phones such as iphone

Permanent challenge

Increasing the performance

One of the possible approaches

• Sequential break down of operations





Micro-processor

- General-purpose PCs
- Servers to desktop, laptop
- Cell-phones such as iphone

Permanent challenge

* Increasing the performance



One of the possible approaches

• Sequential break down of operations



General introduction

Speeding up computations

Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

88



General

Speeding up computations

Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Micro-processor

- General-purpose PCs
- Servers to desktop, laptop
- Cell-phones such as iphone

Permanent challenge

* Increasing the performance

One of the possible approaches

• Sequential break down of operations







Hello, do you remember me.....

General introduction

Speeding up computations Optimizing indexed

communications

break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions





Optimizing indexed communication

Computer networks inspired RCE

• Rapid development of computer networks inspired the advancement of RCE (real time collaborative editors).

n collaborative editing

- Participants are not bound to be in the same location
- Editing shared textual document, to draw a shared graph structure, record ideas during brainstorming meetings.



Documents could be

• Articles, wiki pages and programming source code.

General Introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions



Optimizing indexed communication

Computer networks inspired RCE

• Rapid development of computer networks inspired the advancement of RCE (real time collaborative editors).

In collaborative editing

- Participants are not bound to be in the same location
- Editing shared textual document, to draw a shared graph structure, record ideas during brainstorming meetings.



Documents could be

• Articles, wiki pages and programming source code.

- General ntroduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description
- Conclusions Summary Research directions



Optimizing indexed communication

Computer networks inspired RCE

• Rapid development of computer networks inspired the advancement of RCE (real time collaborative editors).

In collaborative editing

- Participants are not bound to be in the same location
- Editing shared textual document, to draw a shared graph structure, record ideas during brainstorming meetings.



Documents could be

• Articles, wiki pages and programming source code.

- General Introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description
- Conclusions Summary Research directions



Conflicts may occur

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions • In DCE (decentralized collaborative editing), multiple copies of the same document may create confusions.

• To reconcile conflicts, indexing the contents of shared document is considered as basic requirement.

Precision control indexing method (to design)

• Decentralized collaborative editing system





Conflicts may occur

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- In DCE (decentralized collaborative editing), multiple copies of the same document may create confusions.
- To reconcile conflicts, indexing the contents of shared document is considered as basic requirement.

Precision control indexing method (to design)

• Decentralized collaborative editing system





Conflicts may occur

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- In DCE (decentralized collaborative editing), multiple copies of the same document may create confusions.
- To reconcile conflicts, indexing the contents of shared document is considered as basic requirement.

Precision control indexing method (to design)

• Decentralized collaborative editing system





General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations

Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

General introduction

- Speeding up computations
- Optimizing indexed communications

2 Sequential break down of operations

- Problem description
- Contributions
- Decentralized collaborative editing system
 - Problem description
 - Contributions

Conclusions

- Summary
- Research directions



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations

Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions • A 32 bit processor performs operations on 32 bits.

• Any transformation of data structure required decomposition in successive operations on 32 bits.

Old Strategy

Restriction

• To exchange contents of registers, it makes copies of the initial data.

rrors

- More number of registers are required
- Loss of speed/power consumption
- Signal duplication occurs



Restriction

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations
- Problem description
- Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- A 32 bit processor performs operations on 32 bits.
- Any transformation of data structure required decomposition in successive operations on 32 bits.

Old Strategy

• To exchange contents of registers, it makes copies of the initial data.

irrors

- More number of registers are required
- Loss of speed/power consumption
- Signal duplication occurs

イロト イボト イヨト イヨト



General

- introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations
- Problem description
- Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- A 32 bit processor performs operations on 32 bits.
- Any transformation of data structure required decomposition in successive operations on 32 bits.

Old Strategy

Restriction

• To exchange contents of registers, it makes copies of the initial data.

Errors

- More number of registers are required
- Loss of speed/power consumption
- Signal duplication occurs

< ロ > < 同 > < 三 > < 三)



Proposed strategy

in situ design of computation

• It does not require any extra variable other than the variables available as input

xample

break down of operations Problem description

Decentralize collaborative editing

System Problem description Contributions

Conclusions Summary Research directions

• Let $E : \{0, 1\}^2 \longrightarrow \{0, 1\}^2$ be the mapping defined as $E(x_1, x_2) = (x_2, x_3)$

• To compute mapping *E* by a sequence of operations, we have two options:

• A basic program that completes the task.

• An *in situ* program that completes the task.



Proposed strategy

in situ design of computation

• It does not require any extra variable other than the variables available as input

Example

- Sequential break down of operations Problem description
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

• Let $E: \{0, 1\}^2 \longrightarrow \{0, 1\}^2$ be the mapping defined as

$$E(x_1, x_2) = (x_2, x_1)$$

- To compute mapping *E* by a sequence of operations, we have two options:
- A basic program that completes the task.

• An *in situ* program that completes the task.



Proposed strategy

in situ design of computation

• It does not require any extra variable other than the variables available as input

Example

- Sequential break down of operations Problem description
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

• Let $E: \{0, 1\}^2 \longrightarrow \{0, 1\}^2$ be the mapping defined as

$$E(x_1, x_2) = (x_2, x_1)$$

- To compute mapping *E* by a sequence of operations, we have two options:
- A basic program that completes the task.

• An *in situ* program that completes the task.



Comparison

Old

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down o operations

Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

strategy	In situ strategy
$x_0 := x_1$	$x_1 := f_1(x_1, x_2)$
$x_1 := x_2$	$x_2 := f_2(x_1, x_2)$
$x_2 := x_0$	$x_1 := g_1(x_1, x_2)$

$$f_1(x_1, x_2) = f_2(x_1, x_2) = g_1(x_1, x_2) = x_1 \oplus x_2$$

Noticed that

The *in situ* strategy does not make use of an extra variable x_0 which involved in basic strategy.



Comparison

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations

Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Old strategy	In situ strategy
$x_0 := x_1$	$x_1 := f_1(x_1, x_2)$
$x_1 := x_2$	$x_2 := f_2(x_1, x_2)$
$x_2 := x_0$	$x_1 := g_1(x_1, x_2)$

$$f_1(x_1, x_2) = f_2(x_1, x_2) = g_1(x_1, x_2) = x_1 \oplus x_2$$

Noticed that

The *in situ* strategy does not make use of an extra variable x_0 which involved in basic strategy.



How it works

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description
- Problem descriptio
- Decentralize collaborative editing
- Problem descript
- Contributions
- Conclusions Summary Research directions





Interpreting an assignment

General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description

Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions





Interpreting reference value

General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description

Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions





Alternate strategy over fields

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions • IDC strategy to perform *in situ* computation has been proposed by Burckel et al [Bur07].

Alternate Approach

• A strategy based on Bézout's identity

Theorem

Every linear mapping E: $(x, y) \longrightarrow (mx + ny, px + qy)$, where m, n, p, $q \in \mathbb{Z}$, can be computed by a sequence of at most 3 linear assignments with rational coefficients.



Investigating bound for IDC over integers

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations

Problem descripti

Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Objective

• To find the minimum number of assignments required to compute the mapping sequentially using *in situ* strategy.

Theorem

Let

be the mapping over \mathbb{Z}^2 , where F_n is the Fibonacci number. Then the mapping E_{4k+2} is computed with 2k+2 assignments, where $k = 0, 1, 2, \cdots, n$.



Investigating bound for IDC over integers

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations

Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Objective

• To find the minimum number of assignments required to compute the mapping sequentially using *in situ* strategy.

Theorem

Let

$$E_n: (x, y) \longrightarrow (F_{n-1}x + F_ny, F_nx + F_{n+1}y)$$

be the mapping over \mathbb{Z}^2 , where F_n is the Fibonacci number. Then the mapping E_{4k+2} is computed with 2k+2 assignments, where $k = 0, 1, 2, \dots, n$.



Investigating bound for IDC over integers

Theorem

There exists a linear mapping

 $E:(x, y) \longrightarrow (mx + ny, px + qy)$

with m, n, p, $q \in \mathbb{Z}$, that cannot be computed by a sequence of at most 6 linear assignments

$$x := ax + by$$

$$y := cx + dy$$

$$x := ex + fy$$

$$y := gx + hy$$

$$x := ix + jy$$

$$y := kx + ly$$

where a, b, c, d, e, f, g, h, i, j, k, and $I \in \mathbb{Z}$.

< □ ▶ < □ ▶ < □ ▶ < □ ▶ < □ ▶ < □ ▶ 32 / 88

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

collaborative editing system Problem description Contributions

Conclusions Summary Research directions



Identity (Linking coefficients)

Theorem

Let $E : \mathbb{Z}^2 \longrightarrow \mathbb{Z}^2$ be a linear mapping defined as $E(x, y) \longrightarrow (mx+ny, px+qy)$ with integral coefficients $m, n, p, q \in \mathbb{Z}$. Let E be computable by a sequence of linear assignments

$$\begin{aligned} x &:= a_1 x + b_1 y \\ y &:= b_1' x + a_1' y \\ x &:= a_2 x + b_2 y \\ y &:= b_2' x + a_2' y \\ \vdots &:= \vdots \\ x &:= a_k x + b_k y \\ y &:= b_k' x + a_k' y \end{aligned}$$

where $a_i, a_i', b_i, b_i' \in \mathbb{Z}$ are integral coefficients for each i = 1, ..., k. Then the identity

$$mq - pn = a_1 a'_1 \dots a_k a'_k \qquad holds. \tag{1}$$

イロト 不得 トイヨト イヨト

3

33 / 88

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down o operations

Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions



General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description

Decentralized collaborative editing system

Problem description Contributions

Conclusions Summary Research directions

General introduction

- Speeding up computations
- Optimizing indexed communications

Sequential break down of operations

- Problem description
- Contributions

Occentralized collaborative editing system

- Problem description
- Contributions

Conclusions

- Summary
- Research directions



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description

Conclusions Summary Research directions

- To manage conflicting modifications in collaborative editing, indexing method is the fundamental requirement.
- Most recent methods are introduced by

Weiss et al [WUM09]

 Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot-undo: Distributed collaborative editing system on p2p networks.

IEEE Transactions on Parallel and Distributed Systems, 21:1162-1174, 2010.

Preguiça et al [PMSL09]

 Nuno Preguiça and Joan Manuel Marquès and Marc Shapiro and Mihai Leția. A commutative replicated data type for cooperative editing.

A D N A B N A B N A B N

35 / 88

Distributed Computing Systems, International Conference on, 0: 395-403, 2009.



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description

Conclusions Summary Research directions

- To manage conflicting modifications in collaborative editing, indexing method is the fundamental requirement.
- Most recent methods are introduced by

Weiss et al [WUM09]

 Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot-undo: Distributed collaborative editing system on p2p networks.

IEEE Transactions on Parallel and Distributed Systems, 21:1162-1174, 2010.

Preguiça et al [PMSL09]

 Nuno Preguiça and Joan Manuel Marquès and Marc Shapiro and Mihai Leţia. A commutative replicated data type for cooperative editing.

Distributed Computing Systems, International Conference on, 0: 395-403, 2009.


Following situation is not supported

Errors are detected

• These approaches don't support some of the situations.

Identifiers exchange scenario

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions Two characters 'a and 'b' inserted.

- Opdates are executed.
- 3 New character lies b/w 'a' and 'b' ?





Following situation is not supported

Errors are detected

• These approaches don't support some of the situations.

- Problem description

- Two characters 'a' and 'b' inserted.
- Opdates are executed.
- O New character lies b/w 'a' and 'b' ?





Approach by Weiss et al

Order Alteration

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research directio • General form of identifiers corresponding to two insertions are

 $\langle i, s_1, c_1 \rangle$ and $\langle i, s_2, c_2 \rangle$

 $s_1 < s_2$ or $s_1 = s_2$ and $c_1 < c_2$

- Let user U_1 intends to insert new character.
- N = 1. "boundary = 1" then for two identifiers

with p < q (say)

イロト イボト イヨト イヨト



Approach by Weiss et al

Order Alteration

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Summary Research directions • General form of identifiers corresponding to two insertions are

 $\langle i, s_1, c_1 \rangle$ and $\langle i, s_2, c_2 \rangle$

 $s_1 < s_2$ or $s_1 = s_2$ and $c_1 < c_2$

- Let user U_1 intends to insert new character.
- N = 1. "boundary = 1" then for two identifiers

 $p = \langle 0, s, c \rangle$ $q = \langle 0, s, c \rangle,$

with p < q (say)



Approach by Weiss et al

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

• New identifier computed by the algorithm

 $\mathit{id} = \langle 1, \, \mathit{s}, \, \mathit{c}
angle$

• Computed identifier should follows

 $p = \langle 0, \, s, \, c
angle < \langle 1, \, s, \, c
angle < \langle 0, \, s, \, c
angle = q$

∵ 1∉[0, 0]

Order Alteration

Therefore, new id does not lie b/w identifiers p and q .



Order Alteration



Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research directions



• $PosID_a = [1(0:dA)]$ and $PosID_c = [1(0:dB)]$

42 / 88

э



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- Let user B inserts a new character "a₁" between "a" and "c" then algorithm returns new identifier as PosID_{a1} = [1(0 : dA)(1 : dB)]
- If *dA* = (*counter*, *siteID*) = (1, 1) and *dB* = (*counter*, *siteID*) = (1, 2)
- Then $PoslD_a = [1(0:(1, 1))]$ and $PoslD_c = [1(0:(1, 2))]$ • $PoslD_{a_1} = [1(0:(1, 1))(1:(1, 2))]$
- Following the definition (Nuno Preguiça et al.) *PosID_a* < *PosID_{a1}* and *PosID_c* < *PosID_{a1}*

t gives

• $PosID_{a_1} \notin [PosID_a, PosID_c]$



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- Let user B inserts a new character "a₁" between "a" and "c" then algorithm returns new identifier as PosID_{a1} = [1(0 : dA)(1 : dB)]
- If *dA* = (*counter*, *sitelD*) = (1, 1) and *dB* = (*counter*, *sitelD*) = (1, 2)
- Then $PoslD_a = [1(0:(1, 1))]$ and $PoslD_c = [1(0:(1, 2))]$
- $PoslD_{a_1} = [1(0:(1, 1))(1:(1, 2))]$
- Following the definition (Nuno Preguiça et al.)
 PosID_a < PosID_{a1} and PosID_c < PosID_{a1}

t gives

• $PoslD_{a_1} \notin [PoslD_a, PoslD_c]$



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- Let user B inserts a new character "a₁" between "a" and "c" then algorithm returns new identifier as PosID_{a1} = [1(0 : dA)(1 : dB)]
- If *dA* = (*counter*, *siteID*) = (1, 1) and *dB* = (*counter*, *siteID*) = (1, 2)
- Then $PoslD_a = [1(0:(1, 1))]$ and $PoslD_c = [1(0:(1, 2))]$
- $PoslD_{a_1} = [1(0:(1, 1))(1:(1, 2))]$
- Following the definition (Nuno Preguiça et al.)
 PosID_a < PosID_{a1} and PosID_c < PosID_{a1}

It gives

• $PosID_{a_1} \notin [PosID_a, PosID_c].$



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

- Let the shared document is with initial and final points
 Ib =< 0, *NA*, *NA* > and *Ie* =< 10, *NA*, *NA* >.
- User U_1 (with s = 1 and initial clock c = 0) begins





- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- It is possible that, boundary \geq interval/N
- In this case, step=min(interval/N, boundary)=interval/N

It permits that for N = 1, step=9

- Algorithm returns an identifier p =< 9, 1, 1 > such that
 {*lb*, p, *le*} = {< 0, *NA*, *NA* >, < 9, 1, 1 >, < 10, *NA*, *NA* >}
- Let user 1 inserts another character "y" after "x" and an identifier "q" b/w p = < 9, 1, 1 >, le = < 10, NA, NA >
- Let user 2 removes character "x", then it is not sure to perform the desired operation.



- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- It is possible that, boundary \geq interval/N
- In this case, step=min(interval/N, boundary)=interval/N
- It permits that for N = 1, step=9
 - Algorithm returns an identifier p = < 9, 1, 1 > such that • {*lb*, *p*, *le*} = {< 0, *NA*, *NA* >, < 9, 1, 1 >, < 10, *NA*, *NA* >}
 - Let user 1 inserts another character "y" after "x" and an identifier "q" b/w p = < 9, 1, 1 >, le = < 10, NA, NA >
 - Let *user 2* removes character "x", then it is not sure to perform the desired operation.



- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- It is possible that, boundary \geq interval/N
- In this case, step=min(interval/N, boundary)=interval/N
- It permits that for N = 1, step=9
 - Algorithm returns an identifier p = < 9, 1, 1 > such that • {*lb*, *p*, *le*} = {< 0, *NA*, *NA* >, < 9, 1, 1 >, < 10, *NA*, *NA* >}
 - Let user 1 inserts another character "y" after "x" and an identifier "q" b/w p = < 9, 1, 1 >, le = < 10, NA, NA >
 - Let *user 2* removes character "x", then it is not sure to perform the desired operation.



- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- It is possible that, boundary \geq interval/N
- In this case, step=min(interval/N, boundary)=interval/N
- It permits that for N = 1, step=9
 - Algorithm returns an identifier p = < 9, 1, 1 > such that • {*lb*, *p*, *le*} = {< 0, *NA*, *NA* >, < 9, 1, 1 >, < 10, *NA*, *NA* >}
 - Let user 1 inserts another character "y" after "x" and an identifier "q" b/w p = < 9, 1, 1 >, le = < 10, NA, NA >
 - Let user 2 removes character "x", then it is not sure to perform the desired operation.



Errors.....

Infinite Iteration

1... For two identifiers with same values of *i*, Algorithm does not generate any new identifier due to the reason that *interval* = -1.

break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

2... Proposed formula does not work in certain cases

Example

 $p = \langle k_1, 4, 7 \rangle \langle k_2, 9, 5 \rangle$ and $q = \langle k_3, 5, 3 \rangle \langle k_4, 3, 6 \rangle \langle k_5, 3, 9 \rangle$ with assumption BASE = 100, boundary = 10 and $k_1 = k_2 = k_3 = k_4 = k_5 = k$ To insert N lines between two lines identified by p and q,

$$prefix(q, 1) - prefix(p, 1) - 1 = \begin{cases} -1 & \text{if } k_3 = k_1 \\ 0 & \text{if } k_3 = k_1 + 1 \end{cases}$$



Errors.....

Infinite Iteration

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- **1...** For two identifiers with same values of *i*, Algorithm does not generate any new identifier due to the reason that *interval* = -1.
- 2... Proposed formula does not work in certain cases

Example

 $p = \langle k_1, 4, 7 \rangle \langle k_2, 9, 5 \rangle$ and $q = \langle k_3, 5, 3 \rangle \langle k_4, 3, 6 \rangle \langle k_5, 3, 9 \rangle$ with assumption BASE = 100, boundary = 10 and $k_1 = k_2 = k_3 = k_4 = k_5 = k$ To insert N lines between two lines identified by p and q,

$$\mathsf{prefix}(q, 1) - \mathsf{prefix}(p, 1) - 1 = \begin{cases} -1 & \text{if } k_3 = k_1 \\ 0 & \text{if } k_3 = k_1 + 1 \end{cases}$$



Precision control technique

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary



≣ ∽ < C 53 / 88



Contributions

User/site identifiers (USIDs)

Pattern

- $USID \rightarrow 0.00 \dots 0_{p_r+1} d_1 d_2 \dots d_{p_r}$
- where $p_r \geq 1, \ p_\epsilon \geq p_r + \mu, \ \mu \geq 2, \ \mu \in \mathbb{N}$

xample

- Keeping this pattern, we are able to generate 9 × 10^w different epsilons with w = p_r − 1. For p_r ≥ 1
- The minimum value of epsilon is $1 \times 10^{-(p_r+2)}$
- The maximum value is $0.00 \dots 0_{p_r+1}99 \dots 9_{p_r}$
- Property $|USID|_{p_r} = 0$



User/site identifiers (USIDs)

Pattern

- $USID \rightarrow 0.00 \dots 0_{p_r+1} d_1 d_2 \dots d_{p_r}$
- where $p_r \geq 1, \ p_\epsilon \geq p_r + \mu, \ \mu \geq 2, \ \mu \in \mathbb{N}$

Example

break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions 0.001 0.002 0.003 0.004 0.005 0.006 0.007 0.008 0.009

- Keeping this pattern, we are able to generate 9 × 10^w different epsilons with w = p_r − 1. For p_r ≥ 1
- The minimum value of epsilon is $1 \times 10^{-(p_r+2)}$
- The maximum value is $0.00 \dots 0_{p_r+1}99 \dots 9_{p_r}$
- Property $|USID|_{p_r} = 0$



User/site identifiers (USIDs)

Pattern

- $USID \rightarrow 0.00 \dots 0_{p_r+1} d_1 d_2 \dots d_{p_r}$
- where $p_r \geq 1, \ p_\epsilon \geq p_r + \mu, \ \mu \geq 2, \ \mu \in \mathbb{N}$

Example

break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions 0.001 0.002 0.003 0.004 0.005 0.006 0.007 0.008 0.009

- Keeping this pattern, we are able to generate 9 × 10^w different epsilons with w = p_r − 1. For p_r ≥ 1
- The minimum value of epsilon is $1 \times 10^{-(p_r+2)}$
- The maximum value is $0.00 \dots 0_{p_r+1}99 \dots 9_{p_r}$
- Property $|USID|_{p_r} = 0$



Generating character/line identifiers LCIDs

General introduction

Require

- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Summary Research direction

- Points computed for one user must be different from points computed for others
- Set of points computed for each user must be an ordered set
- It should be possible to compute cardinality for the set of points.
- We modified the classical midpoint formula (a + b)/2, for an interval I = [a, b] with 0 ≤ a < b such that ∀ x, y ∈ I

$$f(x, y) = x + \frac{y - x}{2} - \epsilon \quad \text{with } x < y \qquad (2)$$



Generating character/line identifiers LCIDs

General

Require

- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions

- Points computed for one user must be different from points computed for others
- Set of points computed for each user must be an ordered set
- It should be possible to compute cardinality for the set of points.
- We modified the classical midpoint formula (a + b)/2, for an interval I = [a, b] with 0 ≤ a < b such that ∀ x, y ∈ I

$$f(x, y) = x + \frac{y - x}{2} - \epsilon \quad \text{with } x < y \tag{2}$$

イロト イポト イヨト イヨト

58 / 88



Algorithm outlines

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Summary Research direction

- Assume that an interval *I* = [*a*, *b*], 0 ≤ *a* < *b*, presents shared document or *a* and *b* are two *LCIDs* corresponding to two characters/lines.
- To compute new LCIDs, we proceed as follows:

Outlines

Round a and b over the precision p_r
Evaluate ∇ = (b|_{pr}-|a|_{pr}/2) and round it over p_r
If ∇ > USID_{max} then go to the next step
p := ||a|_{pr} + ∇|_{pr} - USID
Verify that p ≤ a and p ≥ b.

イロト イボト イヨト イヨト



Algorithm outlines

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- CONCIUSIONS Summary Research directions

- Assume that an interval *I* = [*a*, *b*], 0 ≤ *a* < *b*, presents shared document or *a* and *b* are two *LCIDs* corresponding to two characters/lines.
- To compute new LCIDs, we proceed as follows:

Outlines

Round a and b over the precision p_r
Evaluate ∇ = (b|p_r-|a|p_r)/2 and round it over p_r
If ∇ > USID_{max} then go to the next step
p := ||a|p_r + ∇|p_r - USID
Verify that p ≤ a and p ≥ b.

A D N A B N A B N A B N



Uniqueness of LCIDs

General

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Summary Research direction Let two participants are assigned with two different USIDs (ϵ_i , ϵ_j such that $\epsilon_i < \epsilon_j$, for 0 < i < j). Then for two operations performed by two participants respectively, two different LCIDs will be generated (over the interval I = [a, b] with $0 \le a < b$).

Example

Theorem

User	Unique Identifiers										
U_1	0.099	0.199	0.299	0.399	0.499	0.599	0.699	0.799	0.899	0.999	
<i>U</i> ₂	0.098	0.198	0.298	0.398	0.498	0.598	0.698	0.798	0.898	0.998	
U ₃	0.097	0.197	0.297	0.397	0.497	0.597	0.697	0.797	0.897	0.997	
U_4	0.096	0.196	0.296	0.396	0.496	0.596	0.696	0.796	0.896	0.996	
U_5	0.095	0.195	0.295	0.395	0.495	0.595	0.695	0.795	0.895	0.995	
U ₆	0.094	0.194	0.294	0.394	0.494	0.594	0.694	0.794	0.894	0.994	



Inserting single character



Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Research directions



≣ ∽্ 62/88



Explanation

General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Contributions

Summary

Research directions



∃ ∽ < ⊂ 63 / 88



Cardinality

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research direction

• Let
$$I = [a, b], 0 \le a < b$$
 and $\gamma = |b - a|$ with $\gamma > 0$.

• Then for the precision p_r , local cardinality that a user can attain is computed as

ormula

.

such that

$$p_r < p_\epsilon < p_d$$



Cardinality

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Summary Research directions • Let $I = [a, b], 0 \le a < b$ and $\gamma = |b - a|$ with $\gamma > 0$.

• Then for the precision p_r , local cardinality that a user can attain is computed as

Formula

۲

$$C_{l} = \begin{cases} \gamma \times 10^{p_{r}} & \text{if } b > a \geq 0, \text{ for } a, b \in \mathbb{N} \\ (\gamma_{1} \times 10^{p_{r}}) - 1 & \text{if } b > a > 0, \text{ for } a, b \in \mathbb{R} \\ & \text{ and } \gamma_{1} = |\gamma|_{p_{r}} \end{cases}$$

such that

$$p_r < p_\epsilon < p_d$$

4 ロ ト 4 日 ト 4 目 ト 4 目 ト 目 の Q () 65 / 88



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research direction

Example

- Suppose that $I_1 = [0, 1]$, $p_r = 3$, $USID_{min} = 0.00001$, and $USID_{max} = 0.0000999$.
- The local cardinality of a set of *LCIDs*, that single user with USID = 0.0000439 can attain, is computed by $C_l = 1 * 10^3 = 1000$.

xample

- If the same user wants to compute local cardinality of *LCIDs* between two *LCIDs* (already computed by him).
- For instance, let $l_2 = [0.7609561, 0.8139561]$ then $\gamma_1 = 0.053$ and $C_l = (0.053 * 10^3) 1 = 52$.



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Example

- Suppose that $I_1 = [0, 1]$, $p_r = 3$, $USID_{min} = 0.00001$, and $USID_{max} = 0.0000999$.
- The local cardinality of a set of *LCIDs*, that single user with USID = 0.0000439 can attain, is computed by $C_l = 1 * 10^3 = 1000$.

Example

- If the same user wants to compute local cardinality of *LCIDs* between two *LCIDs* (already computed by him).
- For instance, let $l_2 = [0.7609561, 0.8139561]$ then $\gamma_1 = 0.053$ and $C_l = (0.053 * 10^3) 1 = 52$.



Assuring order preservation

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Research directions



≣ ৩৭ে 68/88



General ntroduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research direction

Bijection property

Let S' be the set defined over the interval I = [a, b], 0 ≤ a < b, as

$$S' = \{\frac{n'}{10^{p_r}} \mid n' \in \mathbb{N}, a < n' \le \beta, \beta = \gamma * 10^{p_r}\}$$

where γ is the length of the interval I with $|\gamma|_{p_r} > 0$.

• For $1 \le k \le n$, $n \in \mathbb{N}$, let ϵ_k be one of the USIDs then the function $f_k : S' \longrightarrow S''$ defined by $f_k(x') = x' - \epsilon_k, \forall x' \in S'$, is a bijection.



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research directions

Bijection property

Let S' be the set defined over the interval I = [a, b], 0 ≤ a < b, as

$$S' = \{\frac{n'}{10^{p_r}} \mid n' \in \mathbb{N}, a < n' \le \beta, \beta = \gamma * 10^{p_r}\}$$

where γ is the length of the interval I with $|\gamma|_{p_r} > 0$.

• For $1 \le k \le n$, $n \in \mathbb{N}$, let ϵ_k be one of the USIDs then the function $f_k : S' \longrightarrow S''$ defined by $f_k(x') = x' - \epsilon_k, \forall x' \in S'$, is a bijection.



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary

One of the closure properties

• Let S_i , S_j and S_k be the set of *LCIDs* generated for users U_i , U_j and U_k with *USIDs*, ϵ_i , ϵ_j and ϵ_k over the interval $I = [a, b], 0 \le a < b$ respectively.

• Then for any two points $x \in S_i$, $y \in S_j$, x < y and for $\epsilon_j < \epsilon_k$.

$$|x|_{p_r} + \left|\frac{|y|_{p_r} - |x|_{p_r}}{2}\right|_{p_r}\Big|_{p_r} - \epsilon_k \in S_k$$



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary Research direction

One of the closure properties

- Let S_i , S_j and S_k be the set of *LCIDs* generated for users U_i , U_j and U_k with *USIDs*, ϵ_i , ϵ_j and ϵ_k over the interval $I = [a, b], 0 \le a < b$ respectively.
- Then for any two points $x \in S_i$, $y \in S_j$, x < y and for $\epsilon_j < \epsilon_k$.

$$\left||x|_{p_r} + \left|\frac{|y|_{p_r} - |x|_{p_r}}{2}\right|_{p_r}\right|_{p_r} - \epsilon_k \in S_k$$


Analysis of the function

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions Conclusions

Summary

Research directions

- Under rounding conditions.
- **2** Under floating point arithmetic.
- Without rounding conditions,

• Proposed function works nicely.

I	p_e	ϵ_i	ϵ	$Cd(f_c)$	$Cd(f_1)$	$Cd(f_1) - Cd(f_c)$
[0, 0.01]	3	0.003	0.003	3	33	30
[0, 0.01]	4	0.0003	0.0003	33	243	210
[0, 0.01]	5	0.00003	0.00003	257	2216	1959
[0, 0.01]	6	0.000003	0.000003	2049	19380	17331
[0, 0.0001]	7	0.000003	0.0000003	257	2216	1959
[0, 0.0001]	8	0.0000003	0.0000003	2049	19580	17371
[0, 0.0001]	9	0.00000003	0.00000003	32769	165039	132270



Delayed exchange of points

- Users modify/edit shared document one by one and sends updates to all other users.
- It requires to verify the validity of the algorithm for the points including the points that user has received.
- Under this environment, we studied two cases.
 - Random participation
 - Ordered participation

Effect on global cardinality

• This environment may prevent global cardinality in approaching to its upper bound.

ntroduction Speeding up

Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions



Delayed exchange of points

- Users modify/edit shared document one by one and sends updates to all other users.
- It requires to verify the validity of the algorithm for the points including the points that user has received.
- Under this environment, we studied two cases.
 - Random participation
 - Ordered participation

Effect on global cardinality

• This environment may prevent global cardinality in approaching to its upper bound.

General ntroduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

イロト イポト イヨト イヨト



Delayed exchange of points

- Users modify/edit shared document one by one and sends updates to all other users.
- It requires to verify the validity of the algorithm for the points including the points that user has received.
- Under this environment, we studied two cases.
 - Random participation
 - Ordered participation

Effect on global cardinality

• This environment may prevent global cardinality in approaching to its upper bound.

76 / 88

- General introduction
- Speeding up computations Optimizing indexed communications
- Sequential break down of operations Problem description Contributions
- Decentralized collaborative editing system Problem description Contributions
- Conclusions Summary Research directions



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Real time exchange of points

- Contrary to the delayed exchange environment, we propose this environment.
- Under this environment, the identifiers are not required to compute on the basis of remote points and all participants are supposed to be active.

wo situations

- Under this environment, we studied two cases.
 - Random participation
 - Ordered participation



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions

Real time exchange of points

- Contrary to the delayed exchange environment, we propose this environment.
- Under this environment, the identifiers are not required to compute on the basis of remote points and all participants are supposed to be active.

Two situations

- Under this environment, we studied two cases.
 - Random participation
 - Ordered participation



Model (Maximizing global cardinality)

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions



Effect on global cardinality

• This environment makes it possible that global cardinality could reach its upper bound.



Outline

General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions

Summary Research directions

General introduction

- Speeding up computations
- Optimizing indexed communications

2 Sequential break down of operations

- Problem description
- Contributions

Decentralized collaborative editing system

- Problem description
- Contributions

Conclusions

- Summary
- Research directions



We contributed in advancing the capabilities of

- Sequential computation and
- ② Distributed systems

But still a lot of work is to be done both in

- Sequential break-down of operations
- 2 Decentralized collaborative editing system

Dispersion problem

• The removal of point(s) creates the possibility of insertion of point(s).

Each removal does not assure

• The insertion of new point(s).

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions



We contributed in advancing the capabilities of

- Sequential computation and
- Oistributed systems

But still a lot of work is to be done both in

- O Sequential break-down of operations
- Occentralized collaborative editing system

Dispersion problem

• The removal of point(s) creates the possibility of insertion of point(s).

Each removal does not assure

• The insertion of new point(s).

General ntroduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary



We contributed in advancing the capabilities of

- Sequential computation and
- ② Distributed systems

But still a lot of work is to be done both in

- Sequential break-down of operations
- Occentralized collaborative editing system

Dispersion problem

• The removal of point(s) creates the possibility of insertion of point(s).

Each removal does not assure

• The insertion of new point(s).

General ntroduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary

Research directions



We contributed in advancing the capabilities of

- Sequential computation and
- 2 Distributed systems

But still a lot of work is to be done both in

- Sequential break-down of operations
- Occentralized collaborative editing system

Dispersion problem

• The removal of point(s) creates the possibility of insertion of point(s).

Each removal does not assure

• The insertion of new point(s).

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary

Research directions



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

CONCIUSIONS Summary Research directions





Modification/extension for IDC

General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Conclusions Summary Research directions • Combinatorial results and counter examples show that the *in situ* strategies have certain limitations.

Modifications

Limitations

• Modification of the idea is required so that the investigated draw backs could be removed.

Extensions

• Extension of the idea is required to deal with general cases and the real implementation is required.



General introduction

Speeding up computations Optimizing indexed communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description Contributions

Summary

Research directions

Questions?

4 ロト 4 日 ト 4 日 ト 4 日 ト 4 日 ト 2 の 4 C
87 / 88



References

General introduction

Speeding up computations Optimizing indexe communications

Sequential break down of operations Problem description Contributions

Decentralized collaborative editing system Problem description

Conclusions Summary

Research directions

Serge Burckel.

The parallel-sequential duality : Matrices and graphs. *CoRR*, abs/0709.4397, 2007.

Nuno Preguiça, Joan Manuel Marquès, Marc Shapiro, and Mihai Leția.

A commutative replicated data type for cooperative editing.

Distributed Computing Systems, International Conference on, 0:395–403, 2009.

Stéphane Weiss, Pascal Urso, and Pascal Molli.
 Logoot: A scalable optimistic replication algorithm for collaborative editing on p2p networks.
 In *ICDCS*, pages 404–412. IEEE Computer Society, 2009.

イロト イポト イヨト イヨト