## 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
iáa

## Outline

(1) 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
(4) 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

## (1) 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
(4) Conclusions
- Summary
- Research directions


## H. <br> Speeding up computations

## Higher performnace

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



## Speeding up computations

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

## Higher performnace

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


## Significant growth in

(1) Hardware performance
(2) Architecture evolution


## 1 <br> Speeding up computations

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

## Higher performnace

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


## Significant growth in

(1) Hardware performance
(2) Architecture evolution

## Stress is raised on

(1) Compiler technology
(2) Research communities


Optimizing indexed communications

## Sequential

 break down of operationsProblem 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 :'ie 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



## - 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

## 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



## N. 8 <br> Hello, do you remember me...........

General introduction

## Speeding up

 computationsOptimizing indexed communications

Sequential break down of operations
Problem description Contributions

Decentralized collaborative editing system
Problem description Contributions

Conclusions


## Optimizing indexed communication

## Computer networks inspired RCE

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



## 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.


[^0]
## Optimizing indexed communication

General introduction

Sequential break down of operations

Decentralized collaborative editing system

## 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.


## Conflicts may occur

- In DCE (decentralized collaborative editing), multiple copies of the same document may create confusions.



## Conflicts may occur

- 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.



## Conflicts may occur

- 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.
- Decentralized collaborative editing system

Precision control indexing method (to design)


## Outline

## (1) 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
(4) Conclusions
- Summary
- Research directions


## Problem description

## Restriction

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


## Problem description

## Restriction

- 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.



## Problem description

## Restriction

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

General introduction

Speeding up computations
Optimizing indexed communications

Sequential break down of operations
Problem description Contributions

Decentralized collaborative editing
system

## Old Strategy

- 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


## Proposed strategy

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 situ design of computation

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



## Proposed strategy

## in situ design of computation

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


## Example

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

$$
E\left(x_{1}, x_{2}\right)=\left(x_{2}, x_{1}\right)
$$

- To compute mapping $E$ by a sequence of operations, we have two options:



## Proposed strategy

## in situ design of computation

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


## Example

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

$$
E\left(x_{1}, x_{2}\right)=\left(x_{2}, x_{1}\right)
$$

- 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 strategy

$$
\begin{aligned}
x_{0} & :=x_{1} \\
x_{1} & :=x_{2} \\
x_{2} & :=x_{0}
\end{aligned}
$$

## In situ strategy

$$
\begin{aligned}
& x_{1}:=f_{1}\left(x_{1}, x_{2}\right) \\
& x_{2}:=f_{2}\left(x_{1}, x_{2}\right) \\
& x_{1}:=g_{1}\left(x_{1}, x_{2}\right)
\end{aligned}
$$

$$
f_{1}\left(x_{1}, x_{2}\right)=f_{2}\left(x_{1}, x_{2}\right)=g_{1}\left(x_{1}, x_{2}\right)=x_{1} \oplus x_{2}
$$

Noticed that

The in situ strategy does not make use of an extra variable $x_{n}$ which involved in basic strategv.

## Comparison

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 situ strategy

$$
\begin{aligned}
& x_{1}:=f_{1}\left(x_{1}, x_{2}\right) \\
& x_{2}:=f_{2}\left(x_{1}, x_{2}\right) \\
& x_{1}:=g_{1}\left(x_{1}, x_{2}\right)
\end{aligned}
$$

$$
f_{1}\left(x_{1}, x_{2}\right)=f_{2}\left(x_{1}, x_{2}\right)=g_{1}\left(x_{1}, x_{2}\right)=x_{1} \oplus x_{2}
$$

## 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

 introductionSpeeding up computations
Optimizing indexed communications

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


## N. <br> Interpreting an assignment

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


## N. <br> Interpreting reference value

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


## Alternate strategy over fields

- 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 $E:(x, y) \longrightarrow(m x+n y, p x+q y)$, 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

## Objective

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


## General

introduction
Speeding up computations
Optimizing indexed communications

## Sequential

break down of operations

Contributions

## Investigating bound for IDC over integers

General introduction

## Speeding up

 computationsOptimizing incexed communications

Sequential
break down of operations
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

$$
E_{n}:(x, y) \longrightarrow\left(F_{n-1} x+F_{n} y, F_{n} x+F_{n+1} y\right)
$$

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

## Investigating bound for IDC over integers

## Theorem

There exists a linear mapping

## 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

$$
E:(x, y) \longrightarrow(m x+n y, p x+q y)
$$

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

$$
\left.\begin{array}{l}
x:=a x+b y \\
y:=c x+d y \\
x:=e x+f y \\
y:=g x+h y \\
x:=i x+j y \\
y:=k x+l y
\end{array}\right\}
$$

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

## Theorem

Let $E: \mathbb{Z}^{2} \longrightarrow \mathbb{Z}^{2}$ be a linear mapping defined as $E(x, y) \longrightarrow(m x+n y, p x+q y)$

General introduction

## Speeding up

 computationsOptimizing indexed communications

Sequential break down of operations

## Problem description

## Contributions

Decentralized collaborative editing
system
Problem description
Contributions
Conclusions with integral coefficients $m, n, p, q \in \mathbb{Z}$. Let $E$ be computable by a sequence of linear assignments

$$
\left.\left.\begin{array}{c}
x:=a_{1} x+b_{1} y \\
y:=b_{1}^{\prime} x+a_{1}^{\prime} y \\
x:=a_{2} x+b_{2} y \\
y:=b_{2}^{\prime} x+a_{2}^{\prime} y \\
\vdots \\
:=\quad \vdots \\
x:=a_{k} x+b_{k} y \\
y
\end{array}\right\}=b_{k}^{\prime} x+a_{k}^{\prime} y \quad\right\}
$$

where $a_{i}, a_{i}^{\prime}, b_{i}, b_{i}^{\prime} \in \mathbb{Z}$ are integral coefficients for each $i=1, \ldots, k$. Then the identity

$$
\begin{equation*}
m q-p n=a_{1} a_{1}^{\prime} \ldots a_{k} a_{k}^{\prime} \quad \text { holds. } \tag{1}
\end{equation*}
$$

## Outline

## General

 introductionSpeeding 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) 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
(4) Conclusions
- Summary
- Research directions


## Problem description

- 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 p 2 p networks.

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

## Problem description

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

General introduction

Speeding up computations
Optimizing incexed communications

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

## Weiss et al [WUM09]

- Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot-undo: Distributed collaborative editing system on p 2 p 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.


## General

 introductionSpeeding up computations Optimizing indexed communications

Sequential break down of operations
Problem description
Contributions
Decentralized collaborative editing
system
Problem description
Contributions
Conclusions

## Summary

Research dircctions


## Following situation is not supported

## Errors are detected

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

General introduction

Speeding up computations
Optimizing incexed communications

Sequential
break down of operations
Problem description Contributions

Decentralized collaborative editing
system
Problem description Contributions

Conclusions Summary Research directions

## Identifiers exchange scenario

(1) Two characters 'a' and 'b' inserted.
(2) Updates are executed.
(3) New character lies b/w 'a' and 'b' ?


## Approach by Weiss et al

## Order Alteration

- General form of identifiers corresponding to two insertions are

$$
\begin{gathered}
\left\langle i, s_{1}, c_{1}\right\rangle \text { and }\left\langle i, s_{2}, c_{2}\right\rangle \\
s_{1}<s_{2} \text { or } s_{1}=s_{2} \text { and } c_{1}<c_{2}
\end{gathered}
$$

- Let user $U_{1}$ intends to insert new character.
- $N=1$. "boundary $=1$ " then for two identifiers


## Approach by Weiss et al

General introduction

Speeding up computations
Optimizing indexed communications

Sequential
break down of operations

## Order Alteration

- General form of identifiers corresponding to two insertions are

$$
\begin{gathered}
\left\langle i, s_{1}, c_{1}\right\rangle \text { and }\left\langle i, s_{2}, c_{2}\right\rangle \\
s_{1}<s_{2} \text { or } s_{1}=s_{2} \text { and } c_{1}<c_{2}
\end{gathered}
$$

- Let user $U_{1}$ intends to insert new character.
- $N=1$. "boundary $=1$ " then for two identifiers
(1) $p=\langle 0, s, c\rangle$
with $p<q$ (say)
(2) $q=\langle 0, s, c\rangle$,


## Approach by Weiss et al

## Order Alteration

- New identifier computed by the algorithm

$$
i d=\langle 1, s, c\rangle
$$

- Computed identifier should follows

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

$\because 1 \notin[0,0]$
Therefore, new id does not lie $b / w$ identifiers $p$ and $q$.

## Approach by Prejuiça 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 directions


- $\operatorname{Pos} D_{a}=[1(0: d A)]$ and $\operatorname{Pos} D_{c}=[1(0: d B)]$


## Approach by Prejuiça et al

- Let user $B$ inserts a new character " $a_{1}$ " between "a" and " $c$ " then algorithm returns new identifier as $\operatorname{Pos} / D_{a_{1}}=[1(0: d A)(1: d B)]$

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

- If $d A=($ counter, siteID $)=(1,1)$ and $d B=($ counter, siteID $)=(1,2)$


## Approach by Prejuiça et al

- Let user $B$ inserts a new character " $a_{1}$ " between "a" and " $c$ " then algorithm returns new identifier as $\operatorname{Pos} / D_{a_{1}}=[1(0: d A)(1: d B)]$
- If $d A=($ counter, siteID $)=(1,1)$
and $d B=($ counter, sitelD $)=(1,2)$
- Then PosID $D_{a}=[1(0:(1,1))]$ and ${\operatorname{Pos} / D_{c}}=[1(0:(1,2))]$
- PosID $D_{a_{1}}=[1(0:(1,1))(1:(1,2))]$
- Following the definition (Nuno Preguiça et al.) $\operatorname{Pos} D_{a}<\operatorname{Pos}^{2} D_{a_{1}}$ and PosID ${ }_{c}<\operatorname{Pos} D_{a_{1}}$


## Approach by Prejuiça et al

- Let user $B$ inserts a new character " $a_{1}$ " between "a" and " C " then algorithm returns new identifier as

$$
\operatorname{Pos} / D_{a_{1}}=[1(0: d A)(1: d B)]
$$

- If $d A=($ counter, siteID $)=(1,1)$

$$
\text { and } d B=(\text { counter }, \text { sitel } D)=(1,2)
$$

- Then PosID $D_{a}=[1(0:(1,1))]$ and ${\operatorname{Pos} I D_{c}}=[1(0:(1,2))]$
- PosID $a_{a_{1}}=[1(0:(1,1))(1:(1,2))]$
- Following the definition (Nuno Preguiça et al.) Posl $_{a}<\operatorname{Pos}^{\prime} D_{a_{1}}$ and PosID ${ }_{c}<\operatorname{Pos} D_{a_{1}}$

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

## Redundancy of identifiers

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 difections

- Let the shared document is with initial and final points

$$
l b=<0, N A, N A>\text { and } l e=<10, N A, N A>
$$

- User $U_{1}$ (with $s=1$ and initial clock $c=0$ ) begins



## Redundancy of identifiers

- It is possible that, boundary $\geq$ interval/ N


## General

 introductionSpeeding 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 this case, step=min(interval/N, boundary)=interval/ N

It permits that for $N=1$, step $=9$

## 1 <br> Redundancy of identifiers

- 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
- $\{l b, p, l e\}=\{<0, N A, N A\rangle,\langle 9,1,1\rangle,<10, N A, N A\rangle\}$


## Redundancy of identifiers

- 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
- $\{l b, p, l e\}=\{<0, N A, N A\rangle,\langle 9,1,1\rangle,<10, N A, N A\rangle\}$
- Let user 1 inserts another character " $y$ " after " $x$ " and an identifier " $q$ " b/w $p=<9,1,1>$, le $=<10, N A, N A>$


## N. <br> Redundancy of identifiers

- 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
- $\{l b, p, l e\}=\{\langle 0, N A, N A\rangle,\langle 9,1,1\rangle,\langle 10, N A, N A\rangle\}$
- Let user 1 inserts another character " $y$ " after " $x$ " and an identifier " $q$ " b/w $p=<9,1,1>$, le $=<10, N A, N A>$
- Let user 2 removes character " $x$ ", then it is not sure to perform the desired operation.


## Errors.

## General

 introductionSpeeding up computations
Optimizing indexed communications

Sequential break down of operations

## Problem description

Contributions
Decentralized collaborative editing
system
Problem description
Contributions
Conclusions Summary Research directions

## Infinite Iteration

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

## Errors..........

General introduction

Speeding up computations
Optimizing indexed communications

Sequential break down of operations
Problem description Contributions

Decentralized collaborative
editing
system
Problem description Contributions

Conclusions

## Infinite Iteration

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

## Example

$p=\left\langle k_{1}, 4,7\right\rangle\left\langle k_{2}, 9,5\right\rangle$ and $q=\left\langle k_{3}, 5,3\right\rangle\left\langle k_{4}, 3,6\right\rangle\left\langle k_{5}, 3,9\right\rangle$ with assumption $B A S E=100$, 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 $q$,

$$
\operatorname{prefix}(q, 1)-\operatorname{prefix}(p, 1)-1= \begin{cases}-1 & \text { if } k_{3}=k_{1} \\ 0 & \text { if } k_{3}=k_{1}+1\end{cases}
$$

## Precision control technique

## General

 introductionSpeeding up compurations
Optimizing indexed communications

Sequential break down of operations
Problem description
Contributions
Decentralized collaborative
editing
system

Conclusions
Summary
Research directions

-
Basic principle $p_{r}<p_{\epsilon}<p_{d}$

## User/site identifiers (USIDs)

## Pattern

## General

 introductionSpeeding up computations
Optimizing indexed communications

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

- USID $\rightarrow 0.00 \ldots 0_{p_{r}+1} d_{1} d_{2} \ldots d_{p_{r}}$
- where $p_{r} \geq 1, p_{\epsilon} \geq p_{r}+\mu, \mu \geq 2, \mu \in \mathbb{N}$


## 5 <br> User/site identifiers (USIDs)

## Pattern

- USID $\rightarrow 0.00 \ldots 0_{p_{r}+1} d_{1} d_{2} \ldots d_{p_{r}}$
- where $p_{r} \geq 1, p_{\epsilon} \geq p_{r}+\mu, \mu \geq 2, \mu \in \mathbb{N}$

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

| 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 different ensilons with $w_{m}=n_{r}-1$ For $n_{H}>1$
$\square$
- noperty $\mid$ USID|pr $=0$


## User/site identifiers (USIDs)

General introduction
Speeding up computations
Optimizing indexed communications

Sequential
break down of operations
Problem description Contributions

Decentralized collaborative editing system

## Pattern

- USID $\rightarrow 0.00 \ldots 0_{p_{r}+1} d_{1} d_{2} \ldots d_{p_{r}}$
- where $p_{r} \geq 1, p_{\epsilon} \geq p_{r}+\mu, \mu \geq 2, \mu \in \mathbb{N}$


## Example

| 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 $9 \times 10^{w}$ different epsilons with $w=p_{r}-1$. For $p_{r} \geq 1$
- The minimum value of epsilon is $1 \times 10^{-\left(p_{r}+2\right)}$
- The maximum value is $0.00 \ldots 0_{p_{r}+1} 99 \ldots 9_{p_{r}}$
- Property $|U S I D|_{p_{r}}=0$


## Generating character/line identifiers LCIDs

## Require

(1) Points computed for one user must be different from points computed for others
(2) Set of points computed for each user must be an ordered set
(3) It should be possible to compute cardinality for the set of points.


- We modified the classical midpoint formula $(a+b) / 2$, for an interval $I=\lceil a, b\rceil$ with $0<a<b$ such that $\forall x, y \in I$


## Generating character/line identifiers LCIDs

## Require

(1) Points computed for one user must be different from points computed for others
(2) Set of points computed for each user must be an ordered set
(3) 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 $0 \leq a<b$ such that $\forall x, y \in I$

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

## Algorithm outlines

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


## Algorithm outlines

General introduction

Speeding up computations
Optimizing indexed communications

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


## Outlines

(1) Round $a$ and $b$ over the precision $p_{r}$
(2) Evaluate $\nabla=\frac{|b|_{p_{r}}-|a|_{p_{r}}}{2}$ and round it over $p_{r}$
(3) If $\nabla>U S I D_{\max }$ then go to the next step
(9) $p:=\left||a|_{p_{r}}+\nabla\right|_{p_{r}}-U S I D$
(5) Verify that $p \not \leq a$ and $p \nsupseteq b$.

## Uniqueness of LCIDs

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

## Theorem

Let two participants are assigned with two different USIDs ( $\epsilon_{i}$, $\epsilon_{j}$ 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 $0 \leq a<b$ ).

## Example

| 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 |  |  |  |
| $U_{2}$ | 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 |  |  |  |

## N. <br> Inserting single character

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


## Explanation

## General

 introductionSpeeding 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 $I=[a, b], 0 \leq a<b$ and $\gamma=|b-a|$ with $\gamma>0$.


## General

 introduction
## Specing up

 computationsOptimizing indexed
communications
Sequential break down of operations
Problem description
Contributions
Decentralized collaborative
editing
system

Conclusions
Summary
Research directions

- Then for the precision $p_{r}$, local cardinality that a user can attain is computed as


## Cardinality

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

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

- Then for the precision $p_{r}$, local cardinality that a user can attain is computed as


## Formula

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

such that

$$
p_{r}<p_{\epsilon}<p_{d}
$$

## 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 clirections

## Example

- Suppose that $I_{1}=[0,1], p_{r}=3, U S I D_{\min }=0.00001$, and $U S I D_{\max }=0.0000999$.
- The local cardinality of a set of LCIDs, that single user with $U S I D=0.0000439$ can attain, is computed by $C_{I}=1 * 10^{3}=1000$.

General introduction

## Example

- Suppose that $I_{1}=[0,1], p_{r}=3, U S I D_{\min }=0.00001$, and $U S I D_{\max }=0.0000999$.
- The local cardinality of a set of LCIDs, that single user with $U S I D=0.0000439$ can attain, is computed by $C_{I}=1 * 10^{3}=1000$.


## Example

- If the same user wants to compute local cardinality of LCIDs between two LCIDs (already computed by him).
- For instance, let $I_{2}=[0.7609561,0.8139561]$ then

$$
\gamma_{1}=0.053 \text { and } C_{l}=\left(0.053 * 10^{3}\right)-1=52
$$

## N. <br> 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
Summary
Research directions


## Some properties

## General

 introductionSpeeding 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 $S^{\prime}$ be the set defined over the interval $I=[a, b], 0 \leq$ $a<b$, as

$$
S^{\prime}=\left\{\left.\frac{n^{\prime}}{10^{p_{r}}} \right\rvert\, n^{\prime} \in \mathbb{N}, a<n^{\prime} \leq \beta, \beta=\gamma * 10^{p_{r}}\right\}
$$

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

## N. <br> Some properties

General introduction

Speeding up computations
Optimizing indexed communications

Sequential
break down of operations

## Bijection property

- Let $S^{\prime}$ be the set defined over the interval $I=[a, b], 0 \leq$ $a<b$, as

$$
S^{\prime}=\left\{\left.\frac{n^{\prime}}{10^{p_{r}}} \right\rvert\, n^{\prime} \in \mathbb{N}, a<n^{\prime} \leq \beta, \beta=\gamma * 10^{p_{r}}\right\}
$$

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

- For $1 \leq k \leq n, n \in \mathbb{N}$, let $\epsilon_{k}$ be one of the USIDs then the function $f_{k}: S^{\prime} \longrightarrow S^{\prime \prime}$ defined by $f_{k}\left(x^{\prime}\right)=x^{\prime}-\epsilon_{k}, \forall x^{\prime} \in S^{\prime}$, is a bijection.


## N. 1 <br> Some properties

General introduction

Speeding up computations
Optimizing indexed communications

Sequential
break down of operations
Problem description
Contributions
Decentralized collaborative
editing
system

Conclusions

## Summary

Research directions

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 $U S I D s, \epsilon_{i}, \epsilon_{j}$ and $\epsilon_{k}$ over the interval $I=[a, b], 0 \leq a<b$ respectively.


## Some properties

General introduction
 computations
Optimizing indexed communications

## Sequential

break down of operations
Problem description Contributions

Decentralized collaborative

Conclusions

## 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, $\epsilon_{i}, \epsilon_{j}$ and $\epsilon_{k}$ over the interval $I=[a, b], 0 \leq a<b$ respectively.
- Then for any two points $x \in S_{i}, y \in S_{j}, x<y$ and for $\epsilon_{j}<\epsilon_{k}$.

$$
\left||x|_{p_{r}}+\left|\frac{|y|_{p_{r}}-|x|_{p_{r}}}{2}\right|_{p_{r}}\right|_{p_{r}}-\epsilon_{k} \in S_{k}
$$

## 1. <br> 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
(1) Under rounding conditions.
(2) Under floating point arithmetic.
(3) Without rounding conditions,

- Proposed function works nicely.

| I | $p_{e}$ | $\epsilon_{i}$ | $\epsilon$ | $C d\left(f_{c}\right)$ | $\operatorname{Cd}\left(f_{1}\right)$ | $\operatorname{Cd}\left(f_{1}\right)-\operatorname{Cd}\left(f_{c}\right)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $[0,0.01]$ | 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 | 0.0000003 | 0.0000003 | 257 | 2216 | 1959 |
| $[0,0.0001]$ | 8 | 0.00000003 | 0.00000003 | 2049 | 19580 | 17371 |
| $[0,0.0001]$ | 9 | 0.000000003 | 0.000000003 | 32769 | 165039 | 132270 |

## N. <br> Exchange of points

## 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 narticination (2) Ordered participation



## Exchange of points

## 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.
(1) Random participation
(2) Ordered participation



## Exchange of points

## 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.
(1) Random participation
(2) Ordered participation


## Effect on global cardinality

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


## Exchange of points

## 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
- 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.


## Exchange of points

## 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.
(1) Random participation
(2) 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.

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) 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
(4) Conclusions
- Summary
- Research directions


## Further work

## General

 introductionSpeeding up computations
Optimizing indexed communications

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

(1) Sequential computation and
(2) Distributed systems


## Dispersion problem



## Further work

## We contributed in advancing the capabilities of

(1) Sequential computation and
(2) Distributed systems

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

(1) Sequential break-down of operations
(2) Decentralized collaborative editing system

## Further work

## We contributed in advancing the capabilities of

(1) Sequential computation and
(2) Distributed systems

But still a lot of work is to be done both in
(1) 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).

- The insertion of new point(s)


## Further work

## We contributed in advancing the capabilities of

(1) Sequential computation and
(2) Distributed systems

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

(1) 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
Contributions
Decentralized collaborative
editing
system
Problem description
Contributions
Conclusions
Summary
Research directions


## Modification/extension for IDC

## Limitations

- Combinatorial results and counter examples show that the in situ strategies have certain limitations.


## Modifications

- 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

## Questions?

Problem description
Contributions
Decentralized collaborative editing
system
Problem description
Contributions
Conclusions
Summary
Research directions

## References

General introduction Speeding up computations

圊 Serge Burckel．
The parallel－sequential duality：Matrices and graphs．
CoRR，abs／0709．4397， 2007.
嗇 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.
國 Stéphane Weiss，Pascal Urso，and Pascal Molli．
Logoot：A scalable optimistic replication algorithm for collaborative editing on p 2 p networks．
In ICDCS，pages 404－412．IEEE Computer Society， 2009.


[^0]:    - Articles, wiki pages and programming source code

