

Recherche de filtres de Bloom similaires

Application à la recherche par mots clés basée sur une DHT

NDOMBI TSHISUNGU Christian & **DOAN** Cao Sang
Encadrant: M. **MAKPANGOU** Mesaac, Regal

UPMC

2 Mai 2015

Table de contents

- 1 Présentation
- 2 Filtre de Bloom
- 3 Réalisation du travail
- 4 Algorithme des fonctions
 - CREATE_FILTER
 - PUT
 - SEARCH
- 5 Résultat de tests
 - Recherche aléatoire
 - Recherche selective
- 6 Question

Enfin, le texte ! :)

Filtre de Bloom

Algorithme des fonctions

Algorithme des fonctions

CREATE_FILTER

IN : $\sum desc$

FUNCTION : $create_filter(\sum desc)$

OUT : B^{512}

$init(B^{512})$

$x \leftarrow FIRST(\sum desc)$

while $x \neq \emptyset$ **do**

$i \leftarrow SHA_256(x)$

$j \leftarrow i \bmod 512$

$B^{512}[j] \leftarrow 1$

$x \leftarrow NEXT(\sum desc)$

return B^{512}

Algorithme des fonctions

PUT

IN : filtre de Bloom de taille 512 bits B^{512}

FUNCTION : $put(B^{512})$

OUT : \emptyset

$i \leftarrow \text{MAX_LEVEL}$

$\text{vector}_i \leftarrow \text{CREATE_VECTOR}(B^{512}, i)$

$x \leftarrow \text{FIRST}(VA_file)$

while $x \neq \emptyset$ **do**

if $\text{vector}_i = x$ **then**

 BREAK

end

$x \leftarrow \text{NEXT}(VA_file)$

end

Algorithme des fonctions

PUT(suite)

```
if  $vector_i \neq x$  then
   $VA\_file \leftarrow ADD(vector_i)$ 
end
for  $i = MAX\_LEVEL \dots 1$  do
  if  $i = 1$  then
     $vector_i \leftarrow CREATE\_VECTOR(B^{512}, i)$ 
     $CREATE\_FILE(vector_i, B^{512})$ 
  else
     $vector_i \leftarrow CREATE\_VECTOR(B^{512}, i)$ 
     $CREATE\_FILE(vector_i, CREATE\_VECTOR(B^{512}, i - 1))$ 
  end
end
return  $\emptyset$ 
```

Algorithme des fonctions

SEARCH

IN : filtre de Bloom de taille 512 bits B^{512}

FUNCTION : $search(B^{512})$

OUT : $\sum doc$

$i \leftarrow \text{MAX_LEVEL}$

$vector_i \leftarrow \text{CREATE_VECTOR}(B^{512}, i)$

$x \leftarrow \text{FIRST}(VA_file)$

$tmp \leftarrow \text{CREATE_FILE}(i)$

while $x \neq \emptyset$ **do**

if $vector_i \subseteq x$ **then**

$tmp \leftarrow \text{ADD}(x)$

end

$x \leftarrow \text{NEXT}(VA_file)$

end

Algorithme des fonctions

SEARCH(suite)

```
for  $i = MAX\_LEVEL - 1 \dots 1$  do  
   $vector_i \leftarrow CREATE\_VECTOR(B^{512}, i)$   
   $x \leftarrow FIRST(FILE(i + 1))$   
   $tmp \leftarrow CREATE\_FILE(i)$   
  while  $x \neq \emptyset$  do  
    if  $vector_i \subseteq x$  then  
       $tmp \leftarrow ADD(x)$   
    end  
     $x \leftarrow NEXT(FILE(i + 1))$   
  end  
end
```

Algorithme des fonctions

SEARCH(suite)

```
x ← FIRST(FILE(1))
while x ≠ ∅ do
  y ← FIRST(FILE(x))
  while y ≠ ∅ do
    if  $B^{512} \subseteq y$  then
       $\sum doc \leftarrow \text{FIRST}(\text{FILE}(y))$ 
    end
    y ← NEXT(FILE(x))
  end
  x ← NEXT(FILE(1))
end
return  $\sum doc$ 
```


Résultat de tests

Recherche aléatoire

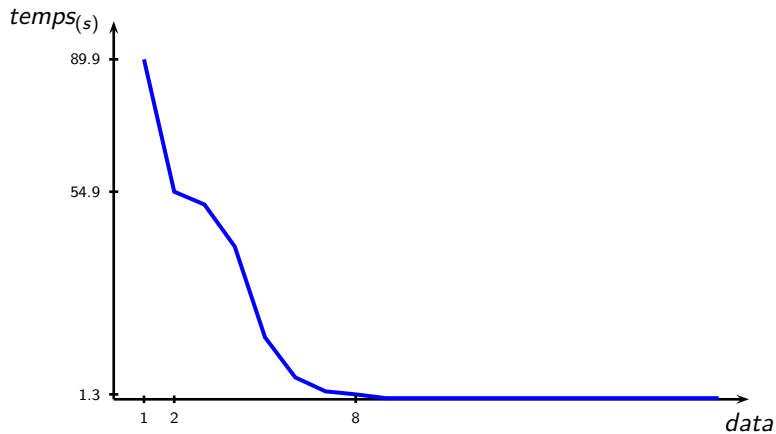


FIGURE: Recherche aléatoire

Résultat de tests

Recherche sélective

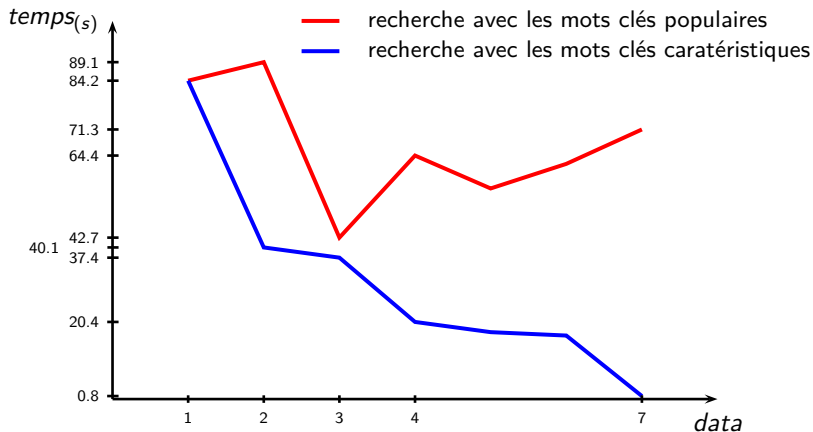


FIGURE: Recherche sélective

Merci de votre attention & Question