

Licenciatura em Engenharia Informática

Algoritmos e Estruturas de Dados

10/01/2022

Trabalho realizado por:

Victor Melo 101099 (50%)

Airton Moreira 100480 (50%)

Introdução

Primeiro trabalho prático de AED , Merkle-Hellman cryptosystem , o objetivo desse trabalho é resolver os problemas de somas de conjunto , dado um determinado conjunto e uma determinada soma , encontrar todos os subconjuntos em que a soma dos elementos desse conjunto seja igual a soma desejada e construir um array de bits em que 1 e 0 fazem corresponder os elementos do conjunto que serão usados para encontrar a soma desejada , 1 significa que o elemento é usado e 0 significa que o elemento não é usado,

Alguns materiais adaptados para implementar as soluções:

- 1. Print sums of all subsets of a given set GeeksforGeeks
- 2. Copying data using the memcpy() function in C educative.io
- 3. <u>Dynamic Memory Allocation in C Geeks for Geeks</u>

Métodos usados para encontrar as soluções

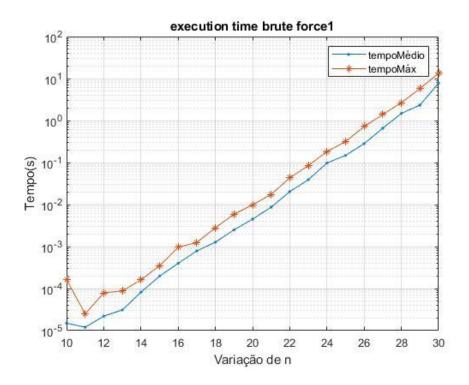
-Bruteforce

Na implementação por força bruta, o array do tipo integer_t *p corresponde ao set de números, onde o problema do subset_sum consiste em encontrar quais números do set p deverão ser utilizados para encontrar a soma desejada (desired_sum).

Na função brute_force, se a soma desejada for igual a soma parcial, popularemos entao o resto do array com zeros (prevenção de erros) e a função brute_force devolve o valor de 1, pois a soma foi encontrada. Já se o índice atual (current_index) for igual a n (tamanho do set p) a função retorna 0, pois já chegou ao fim. Então aplicamos a recursividade pela primeira vez onde chamamos brute_force com o current_index com o valor do próximo indice (current_index + 1), onde rejeitamos o próximo elemento do array de bits no índice atual (que corresponderá a quais indices do set p serão utilizados para obter a soma desejada) e se a variável sol_found, que corresponde se a solução foi encontrada (1) ou não (0), baseada no retorno de brute_force, for verdadeira (1), a funcão brute_force retorna 1, e o elemento b[current_inndex] será 1, pois a solução naquele índice foi encontrada, e por fim retornamos o resultado de brute_force com o indice atual sendo current_index + 1 e a soma parcial sendo a partial_sum (atual) + p[current_index]. Por fim, retornamos o array de bits.

A sua implementação no main com objetivo de otimizar e ganhar velocidade, alocamos logo o espaço em memoria para o integer_t *p (set de números a serem utilizados) com o malloc correspondente ao seu tamanho, após isso preencheremos p com o set fornecido no ficheiro .h correspondente a cada aluno, ou os ficheiros de teste fornecidos pelo Senhor Professor.

Também alocaremos espaço em memória e preencheremos este espaço com zeros para o array de bits integer t *b, onde será preenchido com zeros sem necessidade de um ciclo for.



-Bruteforce otimizado

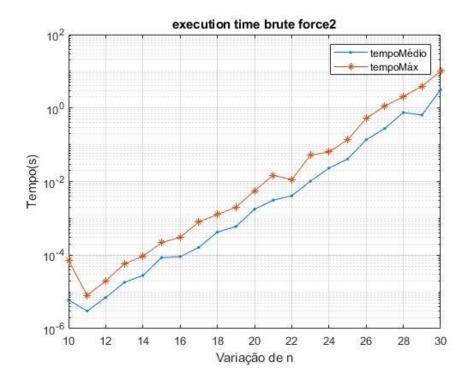
Em Relação ao brute_force1, o brute_force2 (otimizado) consiste em excluir os casos claros onde a soma não terá como ser encontrada, inicialmente excluímos os casos onde a soma parcial é maior que a soma desejada (a funcao retorna 0), após isso quando a soma parcial é maior que a soma desejada, o array de inteiros b será igual a mascara (mask) e a função retorna 1, e a ultima exclusão para otimizar a função, se o índice atual for menor que 0, a execução não tem como trabalhar com índices não existentes no set p, logo a função retorna 0.

Outra diferença é que o integer_t *b neste caso será utilizado como ponteiro para a variável passada, e não é mais um array.

Então se implementará a recursividade, onde se a chamada do brute_force2 com o indice atual = current_index - 1, soma parcial = partial_sum + p[current_index] e a máscara (que inicialmente tem o valor de 0) agora será a operação lógica para extrair o intervalo de bits do current_index, se esta recursão for verdadeira (retorna 1) a função finaliza-se e retorna 1, se não a brute_force2 é chamada recursivamente somente com o índice atual = current_index - 1. Posteriormente a função printBitArray consiste em um loop que irá da int I = 0 até o valor de quantity (inteiro que representa o tamanho do set p utilizado), onde o loop corre o seguinte código: putchar(x & (1u << I)) ? '1''0'), que faz:

Se (1 unsigned, apos feita a operação shift left logical I vezes) & (and logico) x (integer_t de input, utilizado como *b em brute_force2) forem verdadeiros, a função putchar escreve 1, se forem falsos, escreve 0, assim utilizando operações bitwise, o código fica mais rápido ao imprimir os bits que correspondem aos índices do set necessários para obter a soma do subset.

A sua implementação no main com objetivo de otimizar e ganhar velocidade, alocamos logo o espaço em memoria para o integer_t *p (set de números a serem utilizados) com o malloc correspondente ao seu tamanho, após isso preencheremos p com o set fornecido no ficheiro .h correspondente a cada aluno, ou os ficheiros de teste fornecidos pelo Senhor Professor.



- Meet-in-the-middle attack

No meet-in-the-middle o array p é dividido em duas partes, set e set2, que contem todas as somas possíveis de cada um, para dividir o p, usamos o length(p)/2 + 1, para encontrar todas as somas possíveis de cada array usamos a função **subsetSums**, no total são 2**n subconjuntos. para cada elemento, consideramos duas escolhas, incluir no subconjunto e não incluir no subconjunto, os valores encontrados são armazenados nos arrays arrayofallsumsSet1 e arrayofallsumsSet2, que depois são ordenadas usando a função **sort**.

Tendo os arrays arrayofallsumsSet1 e arrayofallsumsSet2 prontos, agora a função m é chamada com os arrays necessários e os seus respetivos tamanhos, o objetivo é encontrar I e j tais que os elementos dos arrays arrayofallsumsSet1[i] e arrayofallsumsSet2[j] somados são iguais a soma desejada, para encontrar o I e o J nos sets originais, é feito a chamada da função brute_force.

A sua implementação no main com objetivo de otimizar e ganhar velocidade, alocamos logo o espaço em memoria para o integer_t *p (set de números a serem utilizados) com o malloc correspondente ao seu tamanho, após isso preencheremos p com o set fornecido no ficheiro .h correspondente a cada aluno, ou os ficheiros de teste fornecidos pelo Senhor Professor.

Também alocaremos espaço em memória e preencheremos este espaço com zeros para o array de bits integer t *b, onde será preenchido com zeros sem necessidade de um ciclo for.

Variação de n

Métodos usados para encontrar os gráficos das soluções

Cada solução gera um ficheiro.txt com o output de sua execução e os bits de cada aluno. Para a obtenção dos gráficos e dos ficheiros de output, limitamos a execução das soluções até o set de número 30. Após obtermos os ficheiros de output, desejamos retirar as informações relativas ao tempo de execução médio, o set que está tal execução e também o tempo médio de execução, para isso, ao corrermos o loop for para cada set desejado, as double somaTempos e tempoMaximo têm valor zero, após isso, no loop para cada soma desejada a double time terá o valor de cpu_time(), e após calcular e fazer o output dos bits a variável time será a diferença de tempo para a realização desses calculos, a função somaTempos tem o valor de time incrementado a ela e se o tempo (time) for maior que o tempoMaximo de execucão, o tempoMaximo será time, agora, de volta ao loop de cada n (set p) o tempo médio será a média da somaTempos / n_sums e após isso, daremos o output do n (i + 10), tempoMedio e tempoMaximo.

Após isso utilizamos um script em bash para obtermos somente os outputs dos dados do tempo de execução, e daí utilizamos um script em matlab para gerarmos os gráficos desejados.

Códigos utilizados:

Brute_force e Brute_force2(otimizado)

```
//
// AED, November 2021
// Solution of the first practical assignement (subset sum problem)
// 101099 Victor Melo 100480 Airton
//
#if __STDC_VERSION__ < 199901L</pre>
#error "This code must must be compiled in c99 mode or later (-std=c99)" // to
handle the unsigned long long data type
#endif
#ifndef STUDENT_H_FILE
#define STUDENT_H_FILE "0000000.h" // "101099.h" "100480.h"
#endif
//
// include files
//
#include <stdio.h>
#include <stdlib.h>
```

```
#include "elapsed_time.h"
#include STUDENT_H_FILE
//
// main program
//
int brute_force(int n, integer_t *p, integer_t desired_sum, int current_index,
integer_t partial_sum, integer_t *b)
{
 if (desired_sum == partial_sum)
 {
   for (int i = current_index; i < n; i++)</pre>
    {
    b[i] = 0;
    }
   return 1;
  }
 if (current_index == n)
  return 0;
  }
 // Nao usar o proximo elemento.(rejeitar o elemento na posicao current_index)
 b[current_index] = 0;
 int sol_found = brute_force(n, p, desired_sum, current_index + 1, partial_sum,
b);
 if (sol_found == 1)
```

```
{
    return 1;
  }
 // Usar o prox elemento
  b[current_index] = 1;
  return brute_force(n, p, desired_sum, current_index + 1, partial_sum +
p[current_index], b);
}
int brute_force2(int n, integer_t *p, int current_index, integer_t partial_sum,
integer_t desired_Sum, integer_t mask, integer_t *b)
{
  if (partial_sum > desired_Sum)
   return 0;
  if (partial_sum == desired_Sum)
  {
    *b = mask;
   return 1;
  }
  if (current_index < 0)</pre>
    return 0;
  if (brute_force2(n, p, current_index - 1, partial_sum + p[current_index],
desired_Sum, (mask | (1 << (current_index))), b))</pre>
    return 1;
```

```
else
   return brute_force2(n, p, current_index - 1, partial_sum, desired_Sum, mask,
b);
}
void printBitArray(integer_t x, int quantity)
{
  for (int i = 0; i <= quantity - 1; i++)</pre>
 {
    putchar(x & (1u << i) ? '1' : '0');</pre>
 }
}
//
// main program
//
int main(void)
{
  int brut = 1; // 1 for brute_force || 2 for brute_force2
 if (brut == 1)
  {
    freopen("brute_force.txt", "w", stdout); // output goes to .txt file
    for (int i = 0; i < n_problems; i++)</pre>
```

```
{
      double somaTempos = 0;
      double tempoMaximo = 0;
      unsigned int n = all_subset_sum_problems[i].n;
      integer_t *p = malloc(n * sizeof(integer_t)); // the weights
      integer_t *b = calloc(n, sizeof(integer_t)); // bit's array -> Use it with
brute_force v1
      p = all_subset_sum_problems[i].p;
     if (n > 30)
      {
      continue;
      }
      for (int j = 0; j < n_sums; j++)</pre>
      {
        double time = cpu_time();
        integer_t desired_sum = all_subset_sum_problems[i].sums[j]; // the
desired sum
        brute_force(n, p, desired_sum, 0, 0, b);
        for (int i = 0; i < n; i++)</pre>
          printf("%lld", b[i]);
        }
```

```
printf("\n");
      time = cpu_time() - time;
      somaTempos += time;
      if (time > tempoMaximo)
        tempoMaximo = time;
      }
    }
    double tempoMedio = somaTempos / n_sums;
    printf("time for: %d %f %f\n", i + 10, tempoMedio, tempoMaximo);
  }
}
if (brut == 2)
{
  freopen("brute_force2.txt", "w", stdout); // output goes to .txt file
  for (int i = 0; i < n_problems; i++)</pre>
  {
    double somaTempos = 0;
    double tempoMaximo = 0;
    unsigned int n = all_subset_sum_problems[i].n;
    integer_t *p = malloc(n * sizeof(integer_t)); // the weights
    p = all_subset_sum_problems[i].p;
```

```
if (n > 30)
 continue;
}
for (int j = 0; j < n_sums; j++)</pre>
{
  double time = cpu_time();
  integer_t desired_sum = all_subset_sum_problems[i].sums[j];
  integer_t x = 0;
 if (brute_force2(n, p, n - 1, 0, desired_sum, 0, &x))
  {
    printBitArray(x, n);
   printf("\n");
  }
  else
  {
    printf("error");
    exit(1);
  }
  time = cpu_time() - time;
  somaTempos += time;
  if (time > tempoMaximo)
  {
    tempoMaximo = time;
  }
}
```

```
double tempoMedio = somaTempos / n_sums;
printf("%d %f %f\n", i + 10, tempoMedio, tempoMaximo);
}

return 0;
}
```

Meet in the Middle

```
//
// AED, November 2021
//
// Solution of the first practical assignement (subset sum problem)
//
// 101099 Victor Melo 100480 Airton
//
#if __STDC_VERSION__ < 199901L
#error "This code must must be compiled in c99 mode or later (-std=c99)" // to handle the unsigned long long data type
#endif
#ifndef STUDENT_H_FILE
#define STUDENT_H_FILE "000000.h"</pre>
```

```
#include <stdio.h>
#include <stdlib.h>
#include "elapsed_time.h"
#include STUDENT_H_FILE
#include <math.h>
#define ARRAYSIZE(a) (sizeof(a)) / (sizeof(a[0]))
#include <string.h>
#include <math.h>
int count;
int brute_force(int n, integer_t *p, integer_t desired_sum, int current_index,
integer_t partial_sum, integer_t *b)
{
   if (desired_sum == partial_sum)
    {
       for (int i = current_index; i < n; i++)</pre>
        {
            b[i] = 0;
        }
        return 1;
    }
```

```
if (current_index == n)
    {
        return 0;
    // Nao usar o proximo elemento.(rejeitar o elemento na posicao current_index)
    b[current_index] = 0;
    int sol_found = brute_force(n, p, desired_sum, current_index + 1,
partial_sum, b);
    if (sol_found == 1)
    {
        return 1;
    }
    // Usar o prox elemento
    b[current_index] = 1;
    return brute_force(n, p, desired_sum, current_index + 1, partial_sum +
p[current_index], b);
}
// Sort the array in ascending order
void sort(integer_t *arr, int length)
{
    integer_t temp = 0;
    for (int i = 0; i < length; i++)</pre>
    {
        for (int j = i + 1; j < length; j++)
        {
            if (arr[i] > arr[j])
```

```
{
                temp = arr[i];
                arr[i] = arr[j];
                arr[j] = temp;
            }
        }
   }
}
/// ALL SUMS OF AN ARRAY
void subsetSums(integer_t *arr, int 1, int r, integer_t sum, int allsums,
integer_t *array)
{
   // Print current subset
   if (1 > r)
   {
       // printf("%d \n", count);
        array[count] = sum;
        count++;
        return;
   }
   // Subset including arr[l]
    subsetSums(arr, l + 1, r, sum + arr[l], allsums, array);
   // Subset excluding arr[l]
```

```
subsetSums(arr, l + 1, r, sum, allsums, array);
}
void m(integer_t *subsets1, integer_t *subsets2, int sub1size, int sub2size,
integer_t desiredsum, integer_t total, int setarraysize, integer_t *set,
integer_t *set2, integer_t *b, integer_t *b1, integer_t *b2)
{
   // printf(" \nDesired sum: %lld ", desiredsum);
   if (desiredsum > total)
    {
        printf(" \n No soluction max desired sum");
        exit(1);
    }
   else
    {
       int test = 0;
        int j = sub1size - 1;
        int i = 0;
        while (test == 0)
        {
           if (subsets1[i] + subsets2[j] == desiredsum)
            {
                // printf("\n The soluction {%lld,%lld}", subsets1[i],
subsets2[j]);
                if (subsets1[i] != 0)
                {
                    brute_force(setarraysize, set, subsets1[i], 0, 0, b1);
                }
```

```
if (subsets2[j] != 0)
                {
                    brute_force(setarraysize, set2, subsets2[j], 0, 0, b2);
                }
                memcpy(b, b1, setarraysize * sizeof(integer_t)); // copy
final_arraysize integer_t's from b1 to b[0]...total[setarraysize]
                memcpy(b + setarraysize, b2, setarraysize * sizeof(integer_t));
                // return b;
                test = 1; // stop the while loop
            }
            if (subsets1[i] + subsets2[j] < desiredsum)</pre>
            {
                i++;
            }
            if (subsets1[i] + subsets2[j] > desiredsum)
            {
                j--;
            }
            if (i > sub1size || j < 0)</pre>
            {
```

```
printf("\nNO soluction ");
               test = 1;
           }
       }
   }
}
/*program to test printPowerSet*/
int main()
{
   freopen("meet_in_the_middle.txt", "w", stdout);
   // fprintf(stderr, "Program configuration:\n");
   // fprintf(stderr, " min_n ..... %d\n", min_n);
   // fprintf(stderr, " max_n ..... %d\n", max_n);
   // fprintf(stderr, " n_sums ..... %d\n", n_sums);
   // fprintf(stderr, " n_problems .. %d\n", n_problems);
   // fprintf(stderr, " integer_t ... %d bits\n", 8 * (int)sizeof(integer_t));
   // printf("\n\t --- MEET IN THE MIDDLE IMPLEMENTATION --- \t\n");
   // set -> sera o array p dado pelo prof do indice 0 ao length(p)/2
   // set2 -> array p do indice lenght(p)/2 ao length(p)
```

```
for (int i = 0; i < n_problems; i++)</pre>
    {
        int n = all_subset_sum_problems[i].n;
        int setarraysize = (n / 2) + 1; // +1 to fix bug
        int allsums = pow(2, setarraysize); // todas as somas possiveis = tamanho
do arrayofallsumsSet1
        int final_arraysize = 2 * setarraysize;
        integer_t set[setarraysize]; // the weights from 0 to length(p)/2
        integer_t set2[setarraysize]; // weights from length(p)/2 to length(p)
        for (int k = 0, j = setarraysize; k < setarraysize; k++, j++)</pre>
        {
            set[k] = all subset sum problems[i].p[k];
            set2[k] = all_subset_sum_problems[i].p[j];
        }
       // Array com todas as somas possiveis
        integer_t arrayofallsumsSet1[allsums];
        integer t arrayofallsumsSet2[allsums];
       // find all sums
        subsetSums(set, 0, setarraysize - 1, 0, allsums, arrayofallsumsSet1);
        count = 0; // make count equal to 0 to avoid bug
        subsetSums(set2, 0, setarraysize - 1, 0, allsums, arrayofallsumsSet2);
        count = 0;
       // sort the arrays
        sort(arrayofallsumsSet1, allsums);
```

```
sort(arrayofallsumsSet2, allsums);
       integer t maxsum = 0;
       for (int i = 0; i < setarraysize; i++)</pre>
       {
           maxsum += set[i];
           maxsum += set2[i];
       }
       double somaTempos = 0;
       double tempoMaximo = 0;
       printf("\n------ Array to work: %d i: %d ------
-\n", n, i);
       for (int j = 0; j < n_sums; j++)</pre>
       {
           double time = cpu_time();
           integer_t desiredsum = all_subset_sum_problems[i].sums[j]; // the
desired sum
           integer_t *b = malloc(final_arraysize * sizeof(integer_t));
           integer_t *b1 = calloc(setarraysize, sizeof(integer_t));
           integer_t *b2 = calloc(setarraysize, sizeof(integer_t));
```

```
// m(arrayofallsumsSet1, arrayofallsumsSet2, allsums, allsums,
desired_sum, maxsum, b);
            m(arrayofallsumsSet1, arrayofallsumsSet2, allsums, allsums,
desiredsum, maxsum, setarraysize, set, set2, b, b1, b2);
            int fixbit;
            if (i % 2 == 0)
            {
               fixbit = 2;
            }
            else
            {
                fixbit = 1;
            }
            printf(" bits: ");
            for (int i = 0; i < (setarraysize * 2) - fixbit; i++) // -fixbits to</pre>
dont count the last bits elements that is trash
            {
                // printf("%lld", b[i]);
                printf("%lld", b[i]);
            }
            printf("\n");
            time = cpu_time() - time;
            somaTempos += time;
            if (time > tempoMaximo)
```

```
free(b);
    free(b1);
    free(b2);
}

// fim do main
    return 0;
}
```

#!/bin/bash

Script em bash para obter dados do tempo de execução

```
input=$1".txt" while IFS= read -r line do echo "$line" | grep " " >>
"times_$1.txt" done < "$input"</pre>
```

Código Matlab

```
ylabel("Tempo(s)");
xlabel("Variação de n")
title('Comparação das Brute Forces')
hold on;
semilogy(n1,tmax1, '*-')
legend("tempoMédio", "tempoMáx", "tempoMédioOTIM", "tempoMáxOTIM");
grid on;
hold off;
```

1000480.h - Output brute_force

12 0.000024 0.000060

13 0.000083 0.000136

14 0.000101 0.000253

15 0.000288 0.000533

16 0.000408 0.000849

17 0.000763 0.001356

18 0.001429 0.002782

19 0.003644 0.006769

20 0.006922 0.010033

1000480.h - Output brute_force2

110000001
1110101101
1111001010
1111001101
1011110100
0001111111
1011100000
0110111101
0100101011
0111010110
1001110000
1101101010
1111110010
0111111000
1001010101
110000011
0001010010
1010001101
1101010101
1101010001
10 0.000004 0.000012
01010000101
10000110001
00011000101
11110000101
10000110000

11 0.000005 0.000021

12 0.000008 0.000025

13 0.000007 0.000029

14 0.000027 0.000113

15 0.000078 0.000217

16 0.000066 0.000278

17 0.000234 0.000576

18 0.000299 0.001146

19 0.001039 0.003059

20 0.002400 0.010678

21 0.005949 0.021119

22 0.011933 0.030249

23 0.033475 0.095809

24 0.061273 0.153618

25 0.090497 0.247674

26 0.171342 0.579786

27 0.202319 0.528323

28 0.436026 1.982954

1000480.h - Output meet_in_the_middle

12 0.000005 0.000007

13 0.000006 0.000007

14 0.000008 0.000011

15 0.000010 0.000013

16 0.000017 0.000042

11101010111100010

17 0.000016 0.000023

18 0.000028 0.000045

19 0.000028 0.000043

20 0.000037 0.000053

101099.h - Output brute_force

	10103	ا ۱۱ ،	- Out	put	DI U	LE_
1011110000						
1110000101						
0010100111						
1111100111						
1011010101						
1101011101						
1111110011						
1101100010						
0001001100						
1100000101						
0110010111						
0000111000						
0001101010						
1001010111						
1000101110						
0111110001						
1001101111						
0001001001						
0111000111						
1011111111						
time for: 10 0.000018 (0.000227					
01011100100						
11101110111						
10010111101						
01110010010						
10111100101						
10100010110						

11110 1011 13 0.000003 1 0.00000

time for: 14 0.000075 0.000159

time for: 15 0.000138 0.000314

time for: 16 0.000413 0.000850

time for: 17 0.000645 0.001294

time for: 18 0.001305 0.002716

time for: 19 0.002589 0.005144

time for: 20 0.004945 0.020835

time for: 21 0.010864 0.020813

time for: 22 0.021285 0.044012

time for: 23 0.041304 0.095063

time for: 24 0.095769 0.202937

time for: 25 0.175305 0.378433

time for: 26 0.380589 0.811496

time for: 27 0.629304 1.581326

1011000101010010000110111010

1101010001101010110010011010

time for: 28 1.585058 3.083862

time for: 30 5.843165 12.156283

101099.h - Output brute_force2

12 0.000008 0.000019

13 0.000011 0.000047

14 0.000035 0.000116

15 0.000056 0.000174

16 0.000087 0.000327

17 0.000272 0.000720

18 0.000403 0.001833

19 0.001056 0.002818

20 0.001445 0.005457

21 0.003056 0.010880

22 0.007950 0.023442

23 0.012019 0.033068

24 0.028580 0.085257

25 0.072658 0.191273

26 0.146703 0.427703

27 0.256804 0.778189

28 0.343930 0.791127

29 1.288769 4.481757

101099.h - Output meet_in_the_middle

10001110100101010101100000

1011101110001100101110001000