MSUP: Model for the Searching of Unidentified People Comparing Binary Vectors Using Jaccard and Dice

José-Sergio Ruiz-Castilla(✉), Farid García-Lamont, José Ángel Regalado-García, Adán Vidal-Peralta, and Carlos Rafael Hernández-Magos

Centro Universitario UAEM Texcoco, Universidad Autónoma del Estado de México, Texcoco, Estado de México, México
jsergioruizc@gmail.com, fglamont@yahoo.com.mx, angel.95rg@gmail.com, adanvidal16@hotmail.com, crhdzmag15@gmail.com

Abstract. The last years in Mexico were reported thousands of missing people. Almost every day were found peoples dead in some place. The authorities open a folder investigation, but most times is not possible an identification effective of the person. In other hand, the familiars report the missing of any member to Public Ministry (PM) or to National Search Commission (NSC). So, the authorities of PM or NSC document the personal data. However, there is not a connection between instances governmental. In this work, we propose a platform and algorithms to find missing people. These algorithms are: The first module is the Characterization of Unidentified People (CUP); the second module is the Characterization of Wanted People (CWP) and finally the Searching for Missing People (SMP). SMP will focus on an algorithm with similarity metrics with the ability to find one or more "similar" people found on the platform. This last module will determine which similar people were found, as well as the dependence on the government where they are physically. To implement the solution, it is necessary to: Establish a vector of characteristics obtained from the CUP module and the CWP module to apply algorithms based on similarity metrics until "matching" and evaluate the proposed algorithms to obtain the best result. For the solution we propose can benefit the institutions that have unidentified corpses under your responsibility. If a human remains is properly characterized, it increases the possibility of identifying and claiming. Therefore, it can reduce or avoid the problem of excess of thousands unclaimed corpses.

Keywords: Similarity metrics · Missing people · People search · Match · Matching

1 Introduction

1.1 Scene in Mexico

When a dead person is found, he usually lacks identification. The elements to identify a person are their physical characteristics. The physical characteristics can be: approximate

L. Martínez-Villaseñor et al. (Eds.): MICAI 2020, LNAI 12469, pp. 337–349, 2020.
https://doi.org/10.1007/978-3-030-60887-3_30

age, weight, height, gender, tattoos, moles, scars, among others. Also, they can be: shoes, clothes, watch, jewelry, etc.

The authorities must lift the corpse under a protocol. The protocol consists of recording the details of the corpse and the environment. An autopsy is performed and said body is protected. Meanwhile, Family members search for the missing person through photographs and physical features.

When a family member disappears, the search begins in public places and at hospitals. The search continues in the morgue or places where corpses are sheltered. However, when a corpse lacks identification, its location is difficult.

This research work seeks to facilitate the search and location of missing and found dead without identifying. The MSUP is proposed (Fig. 2). The MSUP must be implemented in a Web system for governmental instances mainly.

The forensic analysis is very important for identification of corpses. Dorado and Sánchez mentioned in their book ("*What the dead tell*", "*Lo que cuentan los muertos*") as the forensic get all characteristics from corpse or bones. The characteristics are recorded in several forms. The characteristics of the bones and prostheses are useful when a corpse is burned, decomposed or dismembered. The characteristics usually are: approximated age, gender, height, denture, among others [1].

One problem in Mexico is that during the forensic process the data is not completed on the corresponding forms. During autopsy, all data related to the human body must be documented. The main objective is to know the cause of death. However, it is a great opportunity to collect traits for possible identification. It is not a problem only in Mexico, in the work of *Chattopadhyay et al.* documents that Calcutta, India, poses as a disaster the autopsy process in unidentified people. 89% of people studied between 15 to 60 years of age. 7.4% are not recognized before 7 days. Most people were found on roads, paths and rivers. The number of unidentified deaths in the city of Calcutta is quite alarming. It would not be incorrect to describe it as a "*disaster in disguise*" [6].

1.2 Search for Missing People

The "*Registro nacional de datos de personas extraviadas o desaparecidas (RNPED), National data registry of lost or missing persons*", publishes that until 2018 the "Federal Law Statistics" there are 1,171 people missing, while in the "Common Law Statistics" there are, 36,266 people missing or not located. The data recorded are: Date and time of disappearance, country where it disappeared, place of disappearance, nationality, stature, complexion, gender, age, characteristics, ethnicity, disability and place where the disappearance was recorded [2].

The RNPED use the form the Fig. 1 to search to a lost or disappeared person. However, when the person is dead and unidentified is not possible to locate.

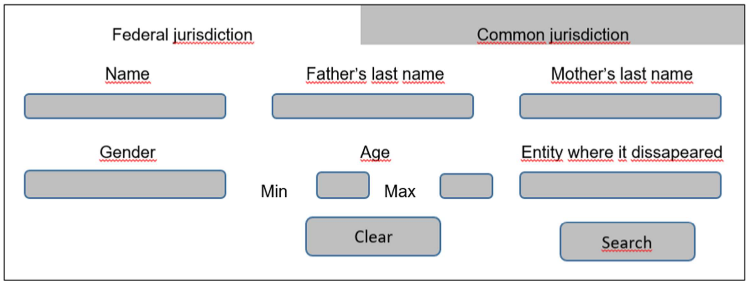


Fig. 1. Form for search people using the website from *RNPED*.

According to the RNPED data, people missing every year from 2014 to 2018 are shown in Table 1. Of which, 26,938 are men and 9,327 are women. In addition, more than 50% are between 15 and 25 years old.

Table 1. Disappeared people from 2014 to 2018 in México [4].

Date	Federal Law	Common Law	Total	Variation
Oct 2014	554	23271	23605	ND
Jun 2015	443	25293	25736	2131
Apr 2015	557	25398	25955	219
Jul 2015	662	25917	26580	625
Oct 2015	916	26670	27586	1006
Jun 2016	946	27215	28161	575
Apr 2016	1027	27162	28189	28
Jul 2016	1044	27428	28472	283
Oct 2016	966	28937	29903	1431
Jun 2017	1030	29912	30942	1039
Total	**8145**	**267203**	**275129**	

2 Related Publications

2.1 People Search

Since 2009 Grandsman proposed to extract blood to children. The above, because the next problem. During the military dictatorship in Argentina (1976–1983), up to 30,000 people disappeared, including an estimated 500 newborn infants and young children who were handed over to military families to be raised as their own. "*Las Abuelas de Plaza de Mayo" (The Grandmothers of the Plaza de Mayo*) is a human rights organization that formed in order to identify their missing grandchildren and reunite them with their biological families. However, the extraction of blood was an illegal practice but can be been an effective strategy if where approved for the government [3].

The cases of unidentified people are not necessarily for homicide. After a disaster, there may be unidentified people such as the 9/11 attacks on the World Trade Center, Hurricane Katrina, or the Southeast Asian tsunami. In the case of 9/11, the DNA was used because the majority of relatives contributed a sample and, on the other hand, the authorities were able to obtain a sample from each unidentified person. [7].

In the work of Andreev et al. exposed a growth of cases of unidentified people in Russia. In this case, these are mostly people who die by drinking alcohol excessively. To identify are used the passport or other documents in the pockets. The 13.5% are unidentified people, but with a growing trend [9].

Bell proposes as a resource to identify people, the dental information. This resource is very effective, however when you do not have the record you lose the possibility. On the other hand, sometimes the bodies are found burned and much dental information is lost. In summary, it is insufficient and other additional characteristics are required. [10]

2.2 Person's Characteristics

The characteristics of a person are related to gender, age, height, weight, skin color, among others. However, there are other more specific ones such as scars, tattoos, moles, etc. Finally, there are other forensic types. Forensic characteristics are the most useful to identify a person. The characteristics of the denture, prosthesis, fractures, absence of limbs.

3 Similarity Metrics

3.1 Binary Similarity

Ie publication Seung-Seok Choi et al., we found 76 formulas for studying binary similarity and distance metrics. In the set of formulas there are formulas that omit (d) that refers to the similarity (0–0). For example: Jaccard, Dice, Czecanowski, 3w-Jaccard, Nei & Li, Sokal & Sneak-I among others [8].

There are formulas for binary coefficient which not use la variable d. This formulas considered for this propose are the next [5] [8].

Jaccard

$$S1(\mathrm{X1}\ \ \mathrm{X1} = \frac{a}{a+b+c+d} \qquad (1$$

Dice

$$S1(\mathrm{X1}\ \ \mathrm{X1}) = \frac{2a}{2a+b+c} \qquad (2$$

3W-Jaccard

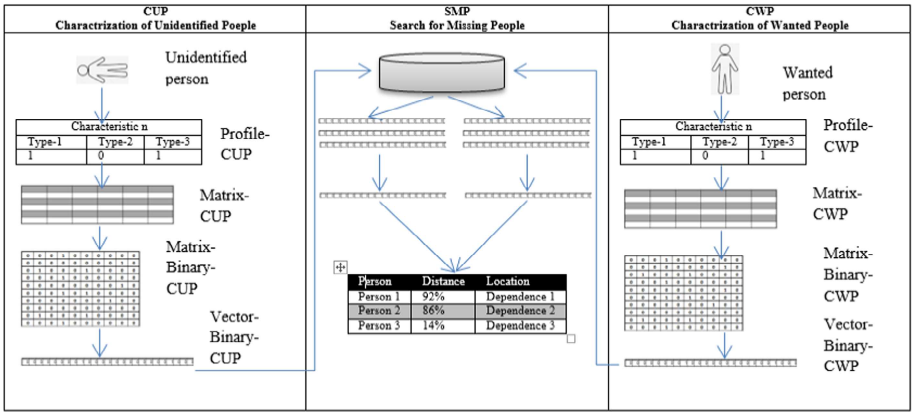
$$S1(\mathrm{X1}\ \ \mathrm{X1}) = \frac{3a}{3a+b+c} \qquad (2$$

We only using solutions with (a), (b), and (c), because the binary matrix has 93% of zeros. Figure 7. With the metrics with (d) the % of similitude was above 90%, the above made very difficult find the best similitude. In this case, Dice and 3 W-Jaccard it assign two or three times the value of (a) but also they omit (d) [5, 8, 11].

4 Method

4.1 MSUP Model

We propose a model for searching the unidentified people. The model not use personal data but the characteristics. We use views for the capture the characteristics. Each matrix or vector has 900 0 s and 1 s. However, each binary vector is stored as a text file independent. The files are read and compared each other. See Fig. 2.



Characteristic n		
Type-1	Type-2	Type-3
1	0	1

Person	Distance	Location
Person 1	92%	Dependence 1
Person 2	86%	Dependence 2
Person 3	14%	Dependence 3

Fig. 2. MSUP: Model for Searching for unidentified people.

The MSUP model is integrated for three modules: The first is *CUP (Characterization of Unidentified People*), the second is *CWP (Characterization for Wanted People*) and the third is *SMP (Search for Missing People*).

The *module CUP* has the function of get the *Profile-CUP* of characteristics of unidentified person. The characteristics are related of her skin, body, hair, clothes, etc. Each characteristic will a binary data as zero or one (0−1). The data binary are recorded in a *Matrix-CUP*. After, The *Matrix-CUP* is converted in at a *Binary-Vector-CUP*; finally, the *Binary-Vector-CUP* is stored in a *Dataset-CUP* in a *Server*.

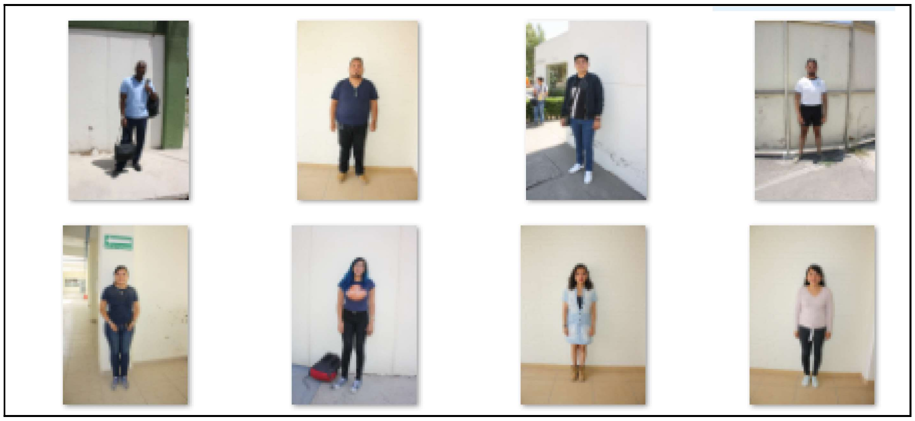


Fig. 3. Photos as source of characteristics.

Multiples photo was taken of volunteer men and women, Fig. 3. These persons were characterized as disappeared persons. With these photos was filled the matrix of characteristics. Accord of matrix of Fig. 4.

The characteristics include of body and clothing and accessories. Other characteristics: the teeth, inlaid teeth, bone prostheses, among others. See the Fig. 4. We obtained about 900 binary elements, shown as 0 s and 1 s.

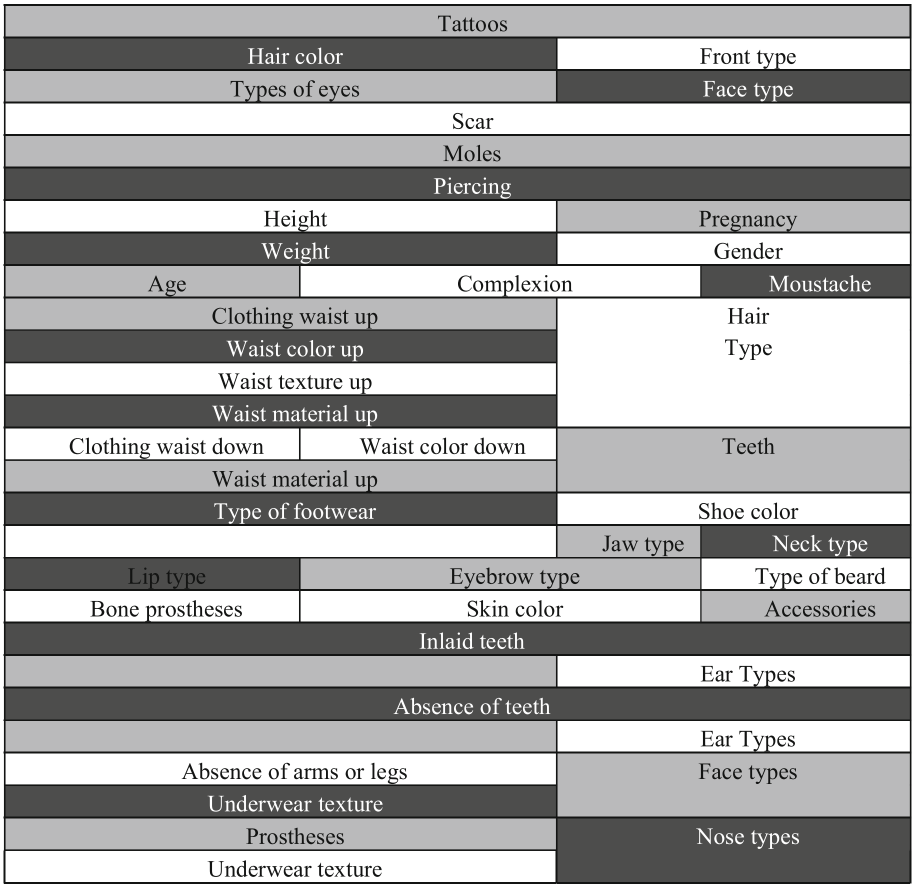
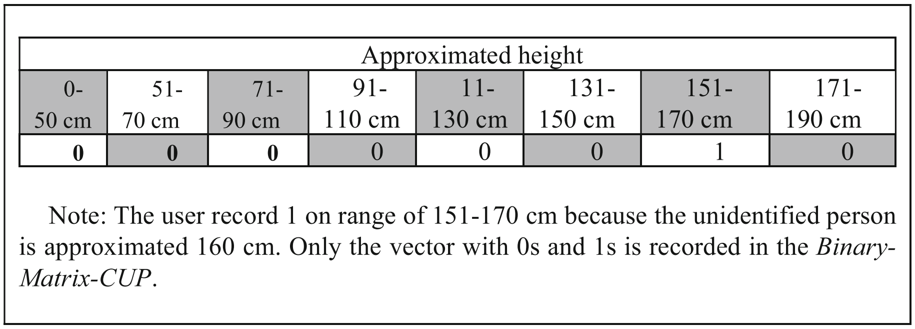

<table>
<tr><td colspan="4">Tattoos</td></tr>
<tr><td colspan="2">Hair color</td><td colspan="2">Front type</td></tr>
<tr><td colspan="2">Types of eyes</td><td colspan="2">Face type</td></tr>
<tr><td colspan="4">Scar</td></tr>
<tr><td colspan="4">Moles</td></tr>
<tr><td colspan="4">Piercing</td></tr>
<tr><td colspan="2">Height</td><td colspan="2">Pregnancy</td></tr>
<tr><td colspan="2">Weight</td><td colspan="2">Gender</td></tr>
<tr><td>Age</td><td colspan="2">Complexion</td><td>Moustache</td></tr>
<tr><td colspan="2">Clothing waist up</td><td colspan="2" rowspan="4">Hair Type</td></tr>
<tr><td colspan="2">Waist color up</td></tr>
<tr><td colspan="2">Waist texture up</td></tr>
<tr><td colspan="2">Waist material up</td></tr>
<tr><td>Clothing waist down</td><td>Waist color down</td><td colspan="2" rowspan="2">Teeth</td></tr>
<tr><td colspan="2">Waist material up</td></tr>
<tr><td colspan="2">Type of footwear</td><td colspan="2">Shoe color</td></tr>
<tr><td colspan="2"></td><td>Jaw type</td><td>Neck type</td></tr>
<tr><td>Lip type</td><td colspan="2">Eyebrow type</td><td>Type of beard</td></tr>
<tr><td>Bone prostheses</td><td colspan="2">Skin color</td><td>Accessories</td></tr>
<tr><td colspan="4">Inlaid teeth</td></tr>
<tr><td colspan="2"></td><td colspan="2">Ear Types</td></tr>
<tr><td colspan="4">Absence of teeth</td></tr>
<tr><td colspan="2"></td><td colspan="2">Ear Types</td></tr>
<tr><td colspan="2">Absence of arms or legs</td><td colspan="2" rowspan="2">Face types</td></tr>
<tr><td colspan="2">Underwear texture</td></tr>
<tr><td colspan="2">Prostheses</td><td colspan="2" rowspan="2">Nose types</td></tr>
<tr><td colspan="2">Underwear texture</td></tr>
</table>

Fig. 4. Set of types of characteristics.

The characteristics included have multiples answers. For each characteristic was created a vector generally for the types. By example, for each tooth was recorded: absence or presence, inlaid teeth end prostheses. In the case of hair: were recorded: presence or absence, color, dyeing color, straight or curly, with extensions, among others.

The process for the characterization is possible using an interface with a set the forms. Each form contains a question about a characteristic. The form has “n” answers, all with zeros. The user only chooses an option, and record 1 as answer. By example, the form shows ages ranges, the user choose the option accord the approximated age of unidentified person. After, the vector is recorded in the *Binary-Vector-CUP*. Each characteristic is processed with the same way. See the Fig. 5.



Approximated height							
0-50 cm	51-70 cm	71-90 cm	91-110 cm	11-130 cm	131-150 cm	151-170 cm	171-190 cm
0	**0**	**0**	0	0	0	1	0

Note: The user record 1 on range of 151-170 cm because the unidentified person is approximated 160 cm. Only the vector with 0s and 1s is recorded in the *Binary-Matrix-CUP*.

Fig. 5. *Binary-Vector* of approximated age.

This method was use for each characteristic. The 0 represents the absence while the 1 represents the presence of the characteristic. This process was used to normalize the data as binary. A vector was created for each characteristic.

The body has been divided into sections. Sections are marked to indicate the existence of a mole, scar, or tattoo during characterization. For now, the tattoo on the body for example a skull is not included.

The *CWP module* has a function of get the *Profile-CWP* of characteristics of wanted persons. The characteristics are of skin, body, clothes, moles, tattoos, etc. The characteristics are recorded as zeros and ones in a *Binary-Vector-CWP*, after in a *Binary-Vector-CWP*.

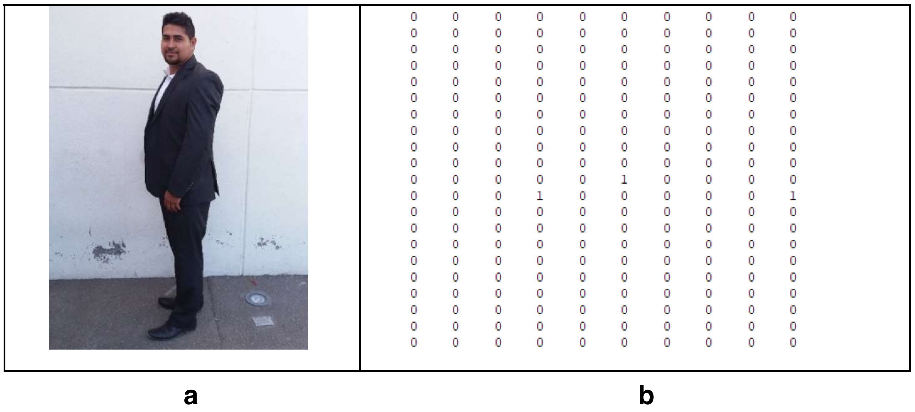
The user that characterizes a wanted person does the following task. The user obtains the characteristics of the person wanted. Then, mark with 1 each characteristic as in Fig. 2. Each binary vector will be added to the *Binary-Matrix-CWP*.

The last module is *SMP* for do match. The "Matching" makes comparison of two strings of characters or 0 s and 1 s. When both strings are equals the distance is 1.0 0r 100%. If there are coincidences then there is a percentage between 0 and 99%. When we obtain a set of records with a percentage each, just choose the highest percentage.

4.2 Characterization of the Missing Person

The characterization is made from a photo or characteristics listed from a member familiar. The physical characteristics and clothes. As well, accessories as like as watch, rings, earrings, and piercings among others. In the Fig. 6 we can see to a) the person's photo and at b) the characterization binary.

The first algorithm is detailed in the Code 1. This code find and open each file for compare two strings end calculate the percentage of similarity. In this case each file have the characterization of a person. The comparing is made with the string of the wanted person and the each string of peoples unidentified registered



```
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  1  0  0  0  0
0  0  0  1  0  0  0  0  0  1
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0
```

Fig. 6. a) The person's photo and at b) The characterization binary.

```
Find the path
Open the folder files
Identification of the folder file according to gender
Set the path
Repeat
Read the file as a characters string
Calculate the similarity of the two strings
Calculate percentage of similarity
Store result
End of repeat
Show the results
```

The second algorithm set the gender. There is a folder for each gender. El gender 1 is male, 2 is female, 3 is transgender and 4 is intersexual. This makes it easier the search because only search in the corresponding folder. See the Code 2.

```
Open the file
Read the file
Generate a binary string
If index 476 = 1 then Gender is Male
If index 477 = 1 then Gender is Female
If index 478 = 1 then Gender is Transgender
If index 479 = 1 then gender is Intersexual
Return Gender
```

The third algorithm allow calculate the grade of similitude. The algorithm make a comparison for find a similitude percentage. For this calculate we use metrics of similitude. See to Code 3.

```
Initialize a = b = c = d = 0
If length (String1) = length(String2)
```

```
Repeat from Index = 0 to length(String2)
If(String[Index]) = 1 and String2[Index] = 1 then a = a+1
If(String[Index]) = 1 and String2[Index] = 0 then b = b+1
If(String[Index]) = 0 and String2[Index] = 1 then c = c+1
If(String[Index]) = 0 and String2[Index] = 0 then d = d+1
Print a,b,c,d
SS3 = a /(a + b + c) //Jaccard
SS2 = (2*a) /((2*a) + b + c) //Dice
SS1 = (3*a) /((3*a) + b + c) //3w-Jaccard
Return SS1, SS2, SS3
```

The Table 2 sample had 41 characterizations of unidentified people, including men and women. Then, two characterizations were made of a man and a woman from the same sample group, but as a wanted person. Matrices and binary vectors were obtained, once the characterizations had been carried out. Finally, we apply the metrics through the corresponding algorithms.

Table 2. Table of binary coefficient. [5, 8, 11]

	1 (presence)	0 (absence)	Sum
1 (presence)	1,1 (a)	1.0 (b)	a + b
0 (absence)	0,1 (c)	0,0 (d)	c + d
	a + c	b + d	A + b + c + d

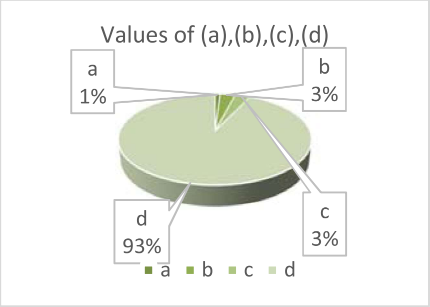


Fig. 7. Ratio of 0’s and 1’s

5 Results

To test the methodology, two cases were studied. The case 1 and 2 yielded the results of Table 3. Are shown the profile of 41 people. Two profiles of unidentified people were included. The results are percentages of similarity. In this case, the higher values indicate more similar to the wanted people.

Table 3. Results of Case 1 and Case 2 with Jaccard, Dice and 3W-Jaccard.

Case 1				Case 2			
	Jaccard	Dice	3 W-Jaccard		Jaccard	Dice	3W-Jaccard
1	0.13	0.22	0.30	1	0.20	0.33	0.42
2	0.25	0.41	0.51	2	0.26	0.42	0.52
3	0.26	0.42	0.52	3	0.17	0.29	0.38
4	0.14	0.24	0.32	4	0.52	0.69	0.77
5	0.17	0.29	0.38	5	0.21	0.35	0.44
6	0.17	0.29	0.38	6	0.24	0.39	0.49
7	0.24	0.38	0.48	7	0.38	0.55	0.65
8	0.18	0.30	0.39	8	0.22	0.37	0.46
9	0.13	0.23	0.31	9	0.06	0.12	0.17
10	0.30	0.46	0.56	10	0.18	0.31	0.40
11	0.30	0.46	0.56	11	0.13	0.22	0.30
12	0.29	0.44	0.55	12	0.06	0.11	0.16
13	0.27	0.43	0.53	13	0.11	0.19	0.26
14	0.25	0.39	0.49	14	0.13	0.23	0.31
15	0.25	0.39	0.49	15	0.12	0.22	0.29
16	0.25	0.41	0.51	16	0.13	0.22	0.30
17	0.67	0.80	0.86	17	0.18	0.31	0.40
18	0.20	0.34	0.43	18	0.19	0.32	0.41
19	0.34	0.51	0.61	19	0.15	0.26	0.34
20	0.24	0.39	0.49	20	0.23	0.37	0.47
21	0.28	0.44	0.54	21	0.18	0.31	0.40
22	0.30	0.46	0.56	22	0.22	0.36	0.46
23	0.41	0.58	0.68	23	0.23	0.37	0.47
24	0.30	0.46	0.56	24	0.19	0.33	0.42
25	0.16	0.27	0.36	25	0.26	0.42	0.52
26	0.21	0.34	0.44	26	0.14	0.24	0.33
27	0.25	0.40	0.50	27	0.09	0.16	0.23
28	0.26	0.42	0.52	28	0.23	0.37	0.47
29	0.28	0.43	0.53	29	0.14	0.25	0.33
30	0.24	0.38	0.48	30	0.16	0.28	0.37
31	0.20	0.33	0.42	31	0.27	0.42	0.52
32	0.31	0.47	0.57	32	0.21	0.34	0.44
33	0.36	0.53	0.62	33	0.18	0.31	0.40

(*continued*)

Table 3. (*continued*)

Case 1				Case 2			
	Jaccard	Dice	3 W-Jaccard		Jaccard	Dice	3W-Jaccard
34	0.27	0.43	0.53	34	0.18	0.30	0.39
35	0.35	0.52	0.62	35	0.10	0.19	0.26
36	0.28	0.43	0.53	36	0.13	0.22	0.30
37	0.33	0.50	0.60	37	0.08	0.14	0.20
38	0.21	0.34	0.44	38	0.25	0.41	0.51
39	0.26	0.42	0.52	39	0.15	0.26	0.34
40	0.14	0.24	0.33	40	0.33	0.50	0.60
41	0.23	0.38	0.48	41	0.20	0.33	0.43

Figure 8 shows case 1. As we can see, the unidentified person 17 corresponds more to the characteristics of the person sought. Also, we can see other cases such as 23, 33, 35 and 37 with higher %, however, it is clear that 17 corresponds to the person sought.

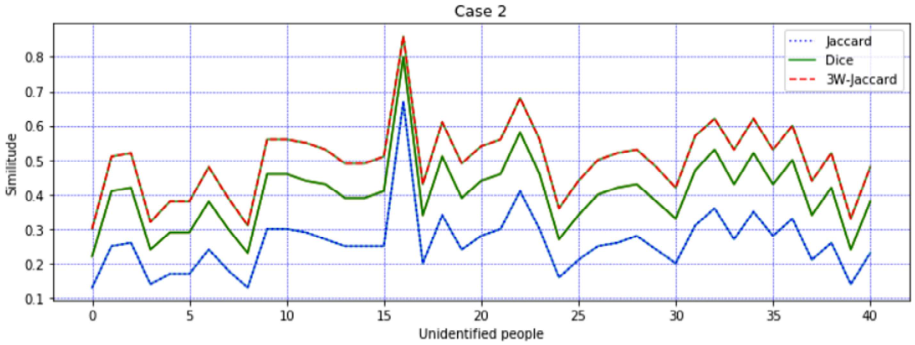


Fig. 8. Graph of case 1, the record 17 is major.

In Case 2, were obtained the results of Table 3. As we can see, record 4 shows the highest %. Similarly, 17 of 41 results. Other high values such as 7, 31, 38 and 40 were found. However, the record 4 corresponds to the person wanted.

In Fig. 9. We can see that register 4 is the highest. Therefore it corresponds to the person wanted. We can see other cases such as 7 or 40. However, the value of register 4 makes it very evident that he is the person wanted.

In this case, of the three techniques used, we can see that the 3 W-Jaccard technique is the most effective, as it shows higher values for all cases. On the other hand, we can see the consistency between the three techniques. It is possible that the values are not 100% due to the characterization of the people. The characterization is done by users through the interfaces and it is possible that they introduce some erroneous data. The above, because most of the characteristics are qualitative.

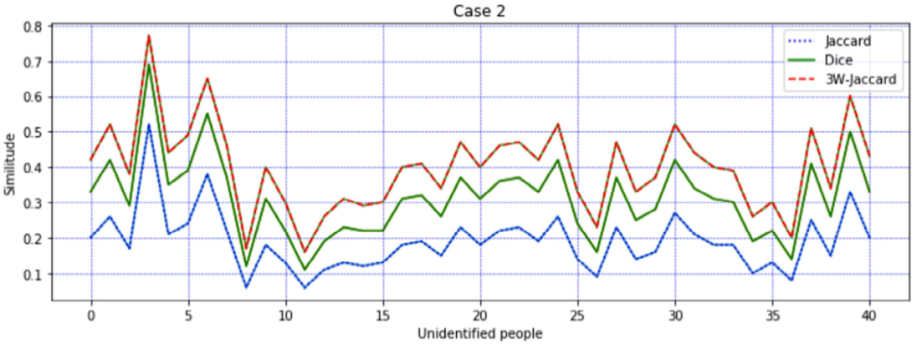


Fig. 9. Graph of Case 2. In this case the record 4 is the major.

6 Conclusions

We can conclude that it is possible to generate the profiles of unidentified people and the profiles of wanted people. With the profiles you can generate matrices and binary vectors in a standard way. Once the profiles are generated they can be stored in text files as strings of 0s and 1s. We can apply similarity metrics and find the closest match. Similar records can be found. In that case, the highest percentages are candidates. Once the most similar record or records have been found, it will be necessary to go to make an identification of the person sought to the government instance. In this case, there are government agencies where the unidentified candidates are to complete the search.

References

1. Dorado E., Sánchez, J.A.: Lo que cuentan los muertos (*What the dead tell).* In: Paidos (ed.) Spain, ISBN: 978-84-998-436-6 (2015)
2. RNPED "Registro Nacional de datos de personas extraviadas o desaparecidas", (NRLMP National Data Registry of Lost or Missing Persons) (2019). www.rndep.segob.gob.mx
3. Grandsman, A.: J. Lat. Am. Carib. Anthropol. **14**(1), 162–184 (2009). https://doi.org/10.1111/j.1935-4940.2009.0001043.x. ISSN 1935-4940. Bay the American Anthropological Association. All rights reserved
4. Centro Diocesano para los Derechos Humanos Fray Juan de Larios. Diagnóstico del Registro Nacional de Datos de Personas Extraviadas o Desaparecidas (RNPED). Saltillo, Coah, México (2017). www.frayjuandelarios.org
5. Rodríguez, S.M.E.: Coeficientes de asociación, Ed. Plaza y Valdez. Ciudad de México (2001)
6. Chattopadhyay, S., Shee, B., Sukul, B.: Unidentified bodies in autopsy–A disaster in disguise. Egypt. J. Forensic Sci. **3**(4), 112–115 (2013). https://doi.org/10.1016/j.ejfs.2013.05.003
7. Ritter N.: Missing Persons and Unidentified Remains: The Nation's Silent Mass Disaster, National Institute of Justice, vol. 256, Washington, DC (2007)
8. Choi, S.-S., Cha, S.-H., Tappert, C.C.: A survey of binary similarity and distance measures. J. Syst. Cybern. Inf. **8**, 43–48 (2010)
9. Andreev, E., Pridemore, W.A., Shkolnikov, V.M., Antonova, O.I.: An investigation of the growing number of deaths of unidentified people in Russia. Eur. J. Pub. Health **18**(3), 252–257 (2008). https://doi.org/10.1093/eurpub/ckm124

10. Bell, G.L.: Dentistry's role in the resolution of missing and unidentified person cases. Dent. Clin. North Am. **45**(2), 293–308 (2001)
11. Batyrshin Ildar Z, Kubysheva Nailya, Solovyev Valery, Villa-Vargas Luis A. Visualization of Similarity Measures for Binary Data and 2 × 2 Tables, Computación y Sistemas, ISSN: 2007-9737, vol 20, 3, pp. 345-353 (2016). https://doi.org/10.13053/cys-20-3-2457