Evaluation – CHC Use Case

CHC Evaluation

- Dataset:
  - Confirmation/Approval required from PRAKSIS
  - **Unsupervised, as of now**
  - Consists of **5 examples**
  - Each **example** consists of **15, 20, 25, 50 and 100 TCN profiles**
  - Each **TCN profile** consists of the following attributes and their domain:
    - **Age –** *[18,120]*
    - **Gender** – *{Male, Female, Other}*
    - **Family** – *{Single Man/Woman, Nuclear, Single Parent Mother/Father, Extended}*
    - **Nationality** – *{<nationality>}*
    - **Religion –** *{<religion>}*
    - **Ethnicity –** *{<ethnicity>}*
    - **Age preference –** *{Don't mind, 18-25, 26-33, 34-43, 44-50, 51-65, 65-120}*
    - **Gender preference** – *{Male, Female, Other, Don't mind}*
    - **Family preference –** *{Single Man/Woman, Nuclear, Single Parent Mother/Father, Extended, Don't mind}*
    - **Nationality preference** – *{Same, Mixed, Don't mind}*
    - **Religion preference** – *{Same, Mixed, Don't mind}*
    - **Ethnicity preference** – *{Same, Mixed, Don't mind}*
    - **Location preference** – *{Don't mind, Ampelokipoi, Menemeni, Kalamaria, Eleftherio-Kordelio, Evosmos, Agios Pavlos, Neapoli, Pefka, Sykies, Nea Efkarpia, Polichni, Stavroupoli, Pylaia, Thessaloniki, Triandria}*
    - **Accessibility preference** – *{Don't mind, Yes, No}*
    - **Rent Period preference** – *{Don't mind, [<date_from>, <date_to>]}*
    - **Share with preference** – *{Don't mind, [<number_minimum>, <number_maximum>]}*

CHC Evaluation

Performance Measures:

1. As of now, **no ground truth** available
    1. **Customer Satisfaction Score** (CSAT) on a Likert scale since system already provides the exact solution. Authors evaluate the BOSS algorithm by comparing its execution time to other exact algorithms such as ODP-IP and ODSS.
        1. We provide the results for the given/approved dataset
        2. Users evaluate the results on a **Likert scale** (5: *Very Satisfied*, 4: *Satisfied*, 3: *Neither Satisfied nor Dissatisfied*, 2: *Dissatisfied*, 1: *Very Dissatisfied)*
        3. In case the **results are not satisfactory**, we try to improve the results based **on users' feedback**. i.e. to be answered by users:
            1. *What makes the results unsatisfactory?*
            2. *How much does it take to compute such results manually? …*
    2. **Silhouette Score** to compute the separation distance between the clusters. Range: [-1,1] where higher score is desired since it means there are small intra-cluster and large inter-cluster average distances.
        1. Nothing is required from users to compute **Silhouette Score**

2. If **ground truth** is **provided**:
    1. **Purity** to compute similarity of two clustering/CS results. Range: [0,1] where 1 means **perfect** match, 0 means **worst** match
    2. **Normalized Mutual Information** (NMI) to compute how much information is shared between a clustering and ground-truth.

        1. Users provide **multiple different examples** and for each of them, they prepare **intuitive grouping results**
        2. We compute the results for the provided examples
        3. We compute **Purity** and **NMI** by comparing **intuitive results vs agents' results**
        4. In case the results for the same example dataset are **not very similar**, we provide the results to users to compute **CSAT (step 1.1.1)**
        5. **Continue step 1.1.2**

CHC Weights

Weights of Preferences
Age Preference: 10
Gender Preference: 9
Family Preference: 5
Nationality Preference: 8
Religion Preference: 6
Ethnicity Preference: 7
Location Preference: 2
Accessibility Preference: 3
Rent Period Preference: 1
Share With Preference: 4

CHC Example_1 → 15 Agents

Agent ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Age	39	37	39	52	70	71	24	47	54	51	71	62	30	63	56
Gender	Female	Male	Male	Other	Other	Other	Male	Other	Male	Female	Other	Male	Other	Male	Female
Family	Extended	Single Man	Single Parent Father	Single Parent Father	Single Woman	Extended	Single Man	Extended	Single Man	Single Woman	Nuclear	Single Man	Nuclear	Single Man	Nuclear
Nationality	nat7	nat1	nat7	nat2	nat5	nat1	nat3	nat4	nat4	nat5	nat7	nat1	nat1	nat2	nat2
Religion	rel4	rel1	rel2	rel7	rel4	rel2	rel6	rel3	rel3	rel7	rel7	rel4	rel3	rel6	rel5
Ethnicity	eth3	eth2	eth3	eth1	eth2	eth3	eth2	eth3	eth6	eth7	eth7	eth5	eth2	eth4	eth5
Age Preference	44-50	34-43	44-50	44-50	34-43/65-120	65-120	18-25/26-33/44-50	44-50	Dont mind	44-50/51-65	18-25/26-33	51-65	34-43	18-25/26-33	Dont mind
Gender Preference	Female	Dont mind	Female	Other	Female	Other	Male	Dont mind	Dont mind	Male	Female	Female	Other	Male	Female
Family Preference	Single Woman/Single Parent Mother	Nuclear	Single Woman/Single Parent Mother	Dont mind	Extended/Nuclear	Single Woman/Single Parent Mother	Extended/Nuclear	Single Parent Mother	Extended	Single Man	Single Woman/Single Parent Mother	Single Woman/Single Parent Mother	Dont mind	Single Parent Mother/Single Parent Father/Extended	Single Parent Mother
Nationality Preference	Mixed	Dont mind	Mixed	Same	Dont mind	Mixed	Mixed	Same	Mixed	Dont mind	Same	Mixed	Same	Dont mind	Dont mind
Religion Preference	Dont mind	Dont mind	Dont mind	Mixed	Dont mind	Mixed	Dont mind	Mixed	Same	Same	Dont mind	Mixed	Dont mind	Dont mind	Dont mind
Ethnicity Preference	Mixed	Dont mind	Mixed	Same	Mixed	Same	Dont mind	Dont mind	Same	Mixed	Mixed	Dont mind	Dont mind	Dont mind	Same
Location Preference	L4/L2/L7	L1	L3	L4	L2/L3	L3/L6	L4	L4/L2/L7	L4/L2/L7	L5	L1/L2/L3	L4	L4/L2/L7	L2/L3	L3/L6
Accessibility Preference	No	Yes	Dont mind	No	Yes	Yes	Dont mind	Dont mind	No	Yes	No	Yes	Yes	Dont mind	Dont mind
Rent Period Preference	1/3/2021-1/7/2022	1/1/2021-1/7/2022	1/1/2021-1/7/2022	1/1/2021-1/4/2021	3/7/2021-12/2/2022	1/1/2021-1/7/2021	1/1/2021-1/7/2022	1/1/2021-1/7/2022	1/5/2021-1/5/2022	15/3/2021-1/4/2022	15/3/2021-1/4/2022	1/5/2021-1/5/2022	1/5/2021-1/5/2022	23/4/2020-17/5/2021	1/9/2021-1/12/2021

Results

***** Not guaranteed that all TCNs will be in a group. Possible that some TCNs can't find a potential group mate because their personal and preferences wouldn't match with others.

* **Singleton**: TCNs who don't match with others.

CHC Example_1 → 15 Agents Result

Total Runtime: 63 seconds
Solution:
Singleton: [8, 1, 4, 11,] --> 4
Group1: [7, 14,] --> 2
Group2: [9, 10, 12, 15,] --> 4
Group3: [2, 13,] --> 2
Group4: [3, 5, 6,] --> 3

CHC Example_2 → 20 Agents Result

Total Runtime: 60 seconds
Solution:
Singleton: [2, 5, 6, 9, 10, 20,] --> 6
Group1: [1, 16,] --> 2
Group2: [15, 17,] --> 2
Group3: [3, 7, 8, 12, 13, 19, 4, 11, 14, 18,] --> 10

CHC Example_3 → 25 Agents Result

Total Runtime: 67 seconds

Solution:
Singleton: [3, 5, 6, 7, 9, 10, 13, 19, 1, 11, 14, 18, 20, 21, 23,] --> 15
Group1: [15, 24,] --> 2
Group2: [2, 8, 17, 16, 25,] --> 5
Group3: [12, 22, 4,] --> 3

CHC Example_4 → 50 Agents Result

Total Runtime: 225 seconds

Solution:
Singleton: [2, 3, 5, 7, 8, 10, 12, 13, 17, 19, 29, 35, 42, 44, 47, 1, 4, 14, 18, 20, 21, 23, 36, 40, 41, 43, 45, 50,] --> 28
Group1: [6, 11, 25, 46,] --> 4
Group2: [9, 24, 38,] --> 3
Group3: [28, 39,] --> 2
Group4: [15, 49,] --> 2
Group5: [27, 37, 16,] --> 3
Group6: [30, 32, 31,] --> 3
Group7: [26, 34,] --> 2
Group8: [22, 33, 48,] --> 3

Results

CHC Example_5 → 100 Agents

Result

Total Runtime: 514 seconds
Solution:
Singleton: [tcn1, tcn3, tcn4, tcn5, tcn6, tcn7, tcn8, tcn11, tcn13, tcn14, tcn15, tcn17, tcn19, tcn20, tcn21, tcn22, tcn23, tcn25, tcn27, tcn28, tcn30, tcn33, tcn34, tcn35, tcn38, tcn41, tcn42, tcn43, tcn44, tcn45, tcn47, tcn48, tcn49, tcn50, tcn52, tcn53, tcn54, tcn55, tcn56, tcn59, tcn60, tcn61, tcn62, tcn63, tcn64, tcn65, tcn66, tcn67, tcn68, tcn69, tcn70, tcn72, tcn73, tcn74, tcn75, tcn77, tcn78, tcn79, tcn80, tcn81, tcn82, tcn83, tcn84, tcn85, tcn89, tcn95, tcn96, tcn100,] → 68
Group1: [tcn16, tcn31,] → 2
Group2: [tcn24, tcn46,] → 2
Group3: [tcn10, tcn12,] → 2
Group4: [tcn9, tcn18,] → 2
Group5: [tcn36, tcn71,] → 2
Group6: [tcn29, tcn40,] → 2
Group7: [tcn58, tcn93,] → 2
Group8: [tcn26, tcn98,] → 2
Group9: [tcn2, tcn88, tcn99,] → 3
Group10: [tcn39, tcn57,] → 2
Group11: [tcn32, tcn37, tcn51, tcn76, tcn86, tcn87, tcn94, tcn97,] → 8
Group12: size - 3, [tcn90, tcn91, tcn92,]