Detecting Tourist's Preferences by Sentiment Analysis in Smart Cities

Zahra Abbasi-Moud
Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
zahra.abbasi@birjand.ac.ir

Hamed Vahdat-Nejad
Faculty of Electrical and Computer Engineering
University of Birjand
Birjand, Iran
vahdatnejad@birjand.ac.ir

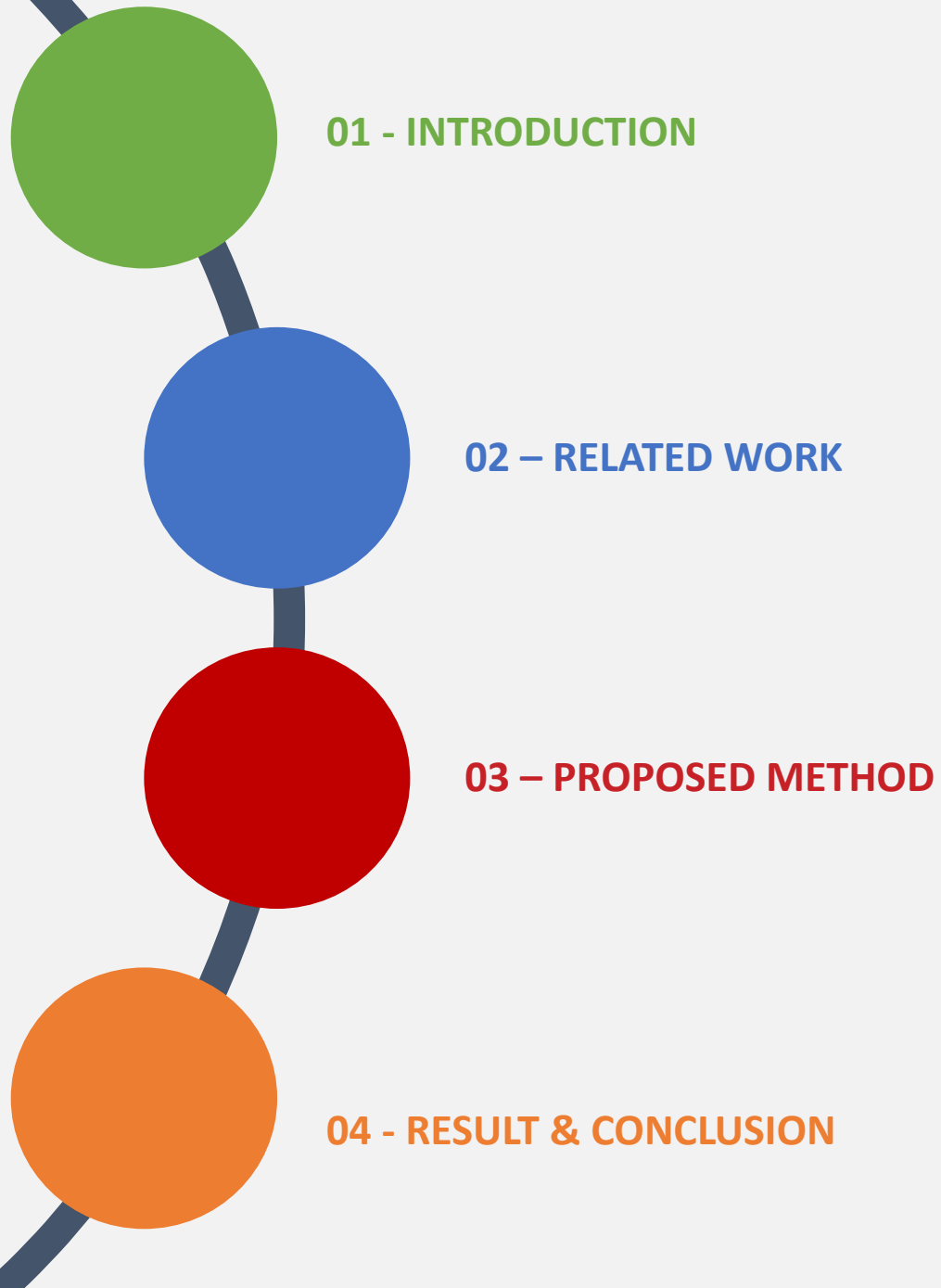
Wathiq Mansoor
Department of Electrical Engineering
University of Dubai
Dubai, United Arab Emirates
wmansoor@ud.ac.ae

IEEE Global Conference On Internet of Things(GCIoT)

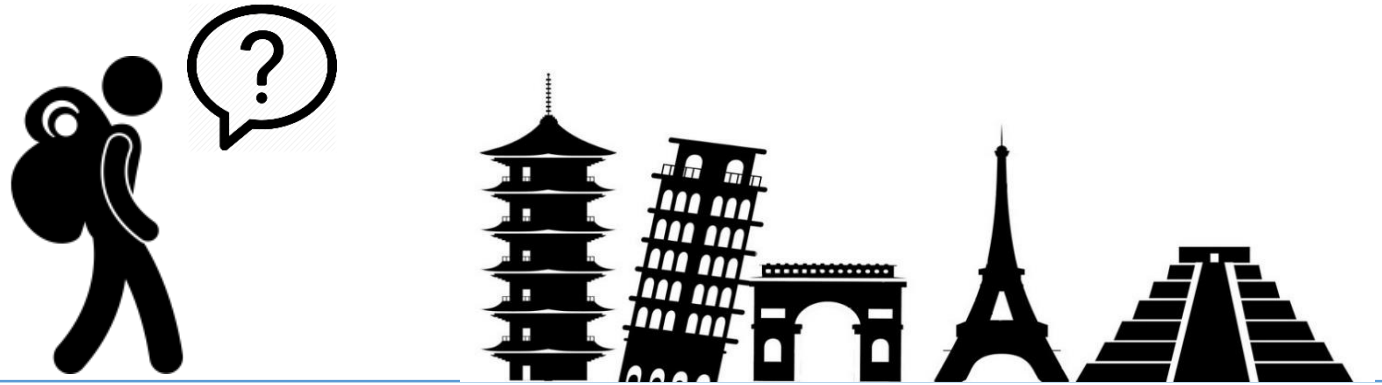
Dubai, United Arab Emirates

Dec 2019

TABLE OF CONTENTS

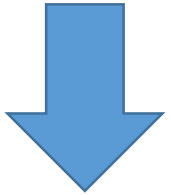


Extracting tourists' preferences is vital for delivering personalized services.



Research problem: How to extract tourist preferences?

Methods of extracting tourist preferences	Approach	Characteristic	Gap
	Based on tourist visit records	Inflexible	The user might become unhappy after the visit
	Data mining on tourist reviews	Flexible	Can not distinguish priorities from disgusts

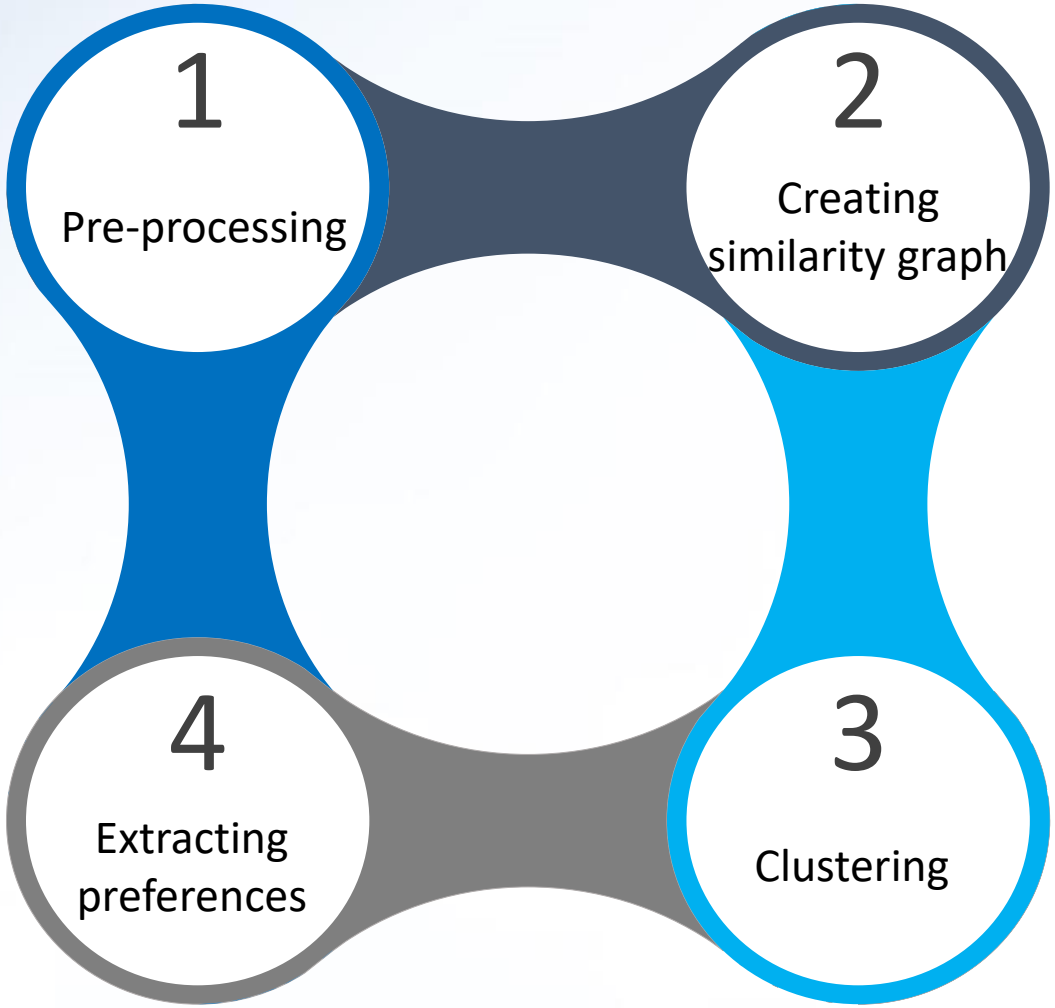


Solution: Sentiment analysis

Use of text processing to identify affective and sensual states of a text.

Proposed method

Tourist reviews



Proposed method:

Step 1:
Pre-processing

1. Part of speech tagging

pronoun	verb	verb	noun
I	love	visiting	museums.

2. Stop words elimination

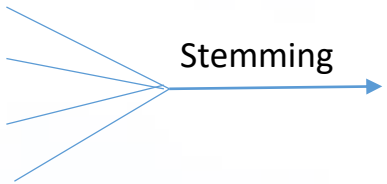
Stop words are frequent words that do not give us specific information. Example: "a“, "and“, "but“, “and "

3. Extracting nouns

I love visiting museums. ⟹ museum is extracted.

4. Stemming

Consult
Consulting
Consultant
Consultants



Consult

Proposed method:

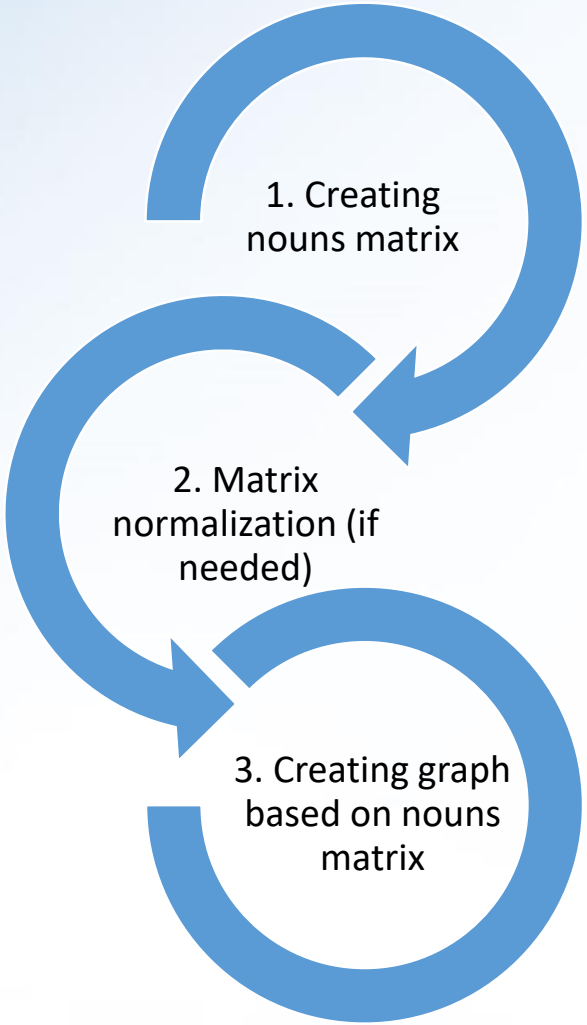
Step 2:
Creating similarity graph

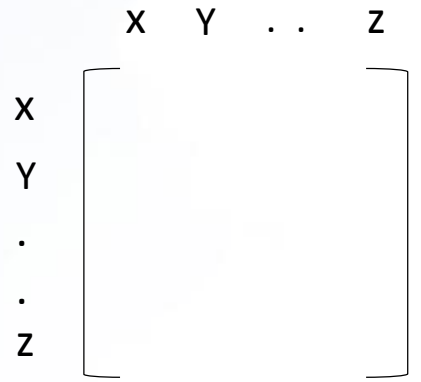
Similarity measures:

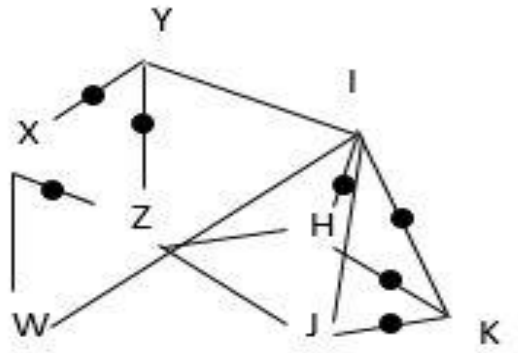
1- Wu-Palmer

2- Extended Gloss Overlaps

3- Hybrid







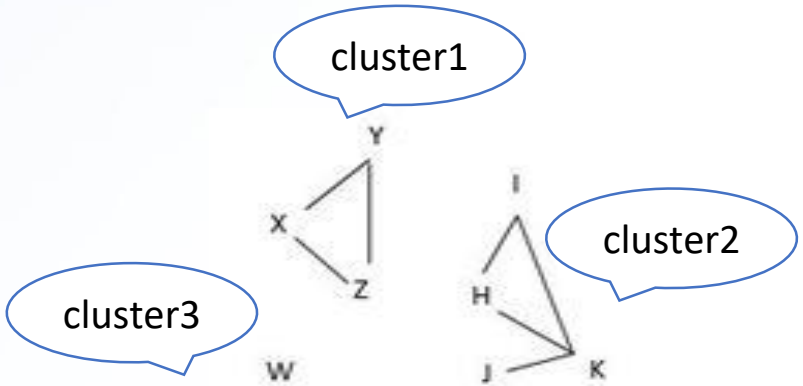
- Edges that weigh more than the threshold

  Threshold = 0.8

Proposed method:

Step 3:
Clustering

1. Removing unnecessary edges



2. Transferring sentences to the clusters that include any of their words

- Sentence1: “azb” → Cluster1
- Sentence2: “jkx” → Cluster1 , Cluster2

Proposed method:

Step 4: Extracting preferences

1. Scoring clusters
2. Returning the highest scored cluster

$$Score(cluster_i) = TF(cluster_i) \times Score_{Sentiment\ Analysis}(cluster_i)$$

$TF(cluster_i) =$ How many times the words of $cluster_i$ are appeared in the user's reviews?

$$Score_{Sentiment\ Analysis}(cluster_i) = \frac{\sum Sentiment\ analysis\ score\ of\ each\ sentence\ of\ (cluster_i)}{Total\ number\ of\ sentences\ in\ (cluster_i)}$$

✓ The cluster with the highest score represents the user's preferences.

Dataset:

Data source

www.TripAdvisor.com

Time period

January to July 2018

Number of users

100

Number of comments

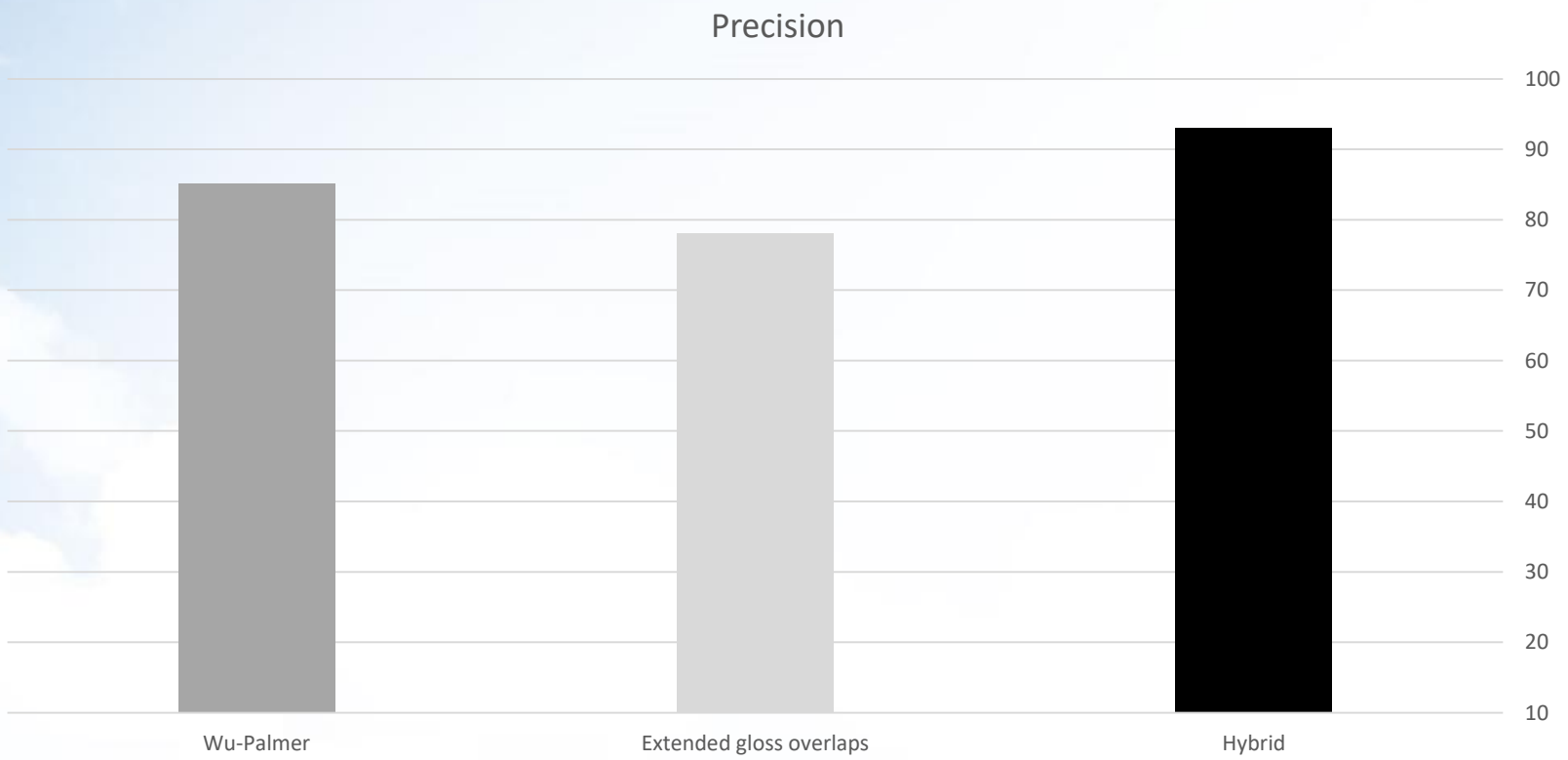
1286

Number of sentences

3332

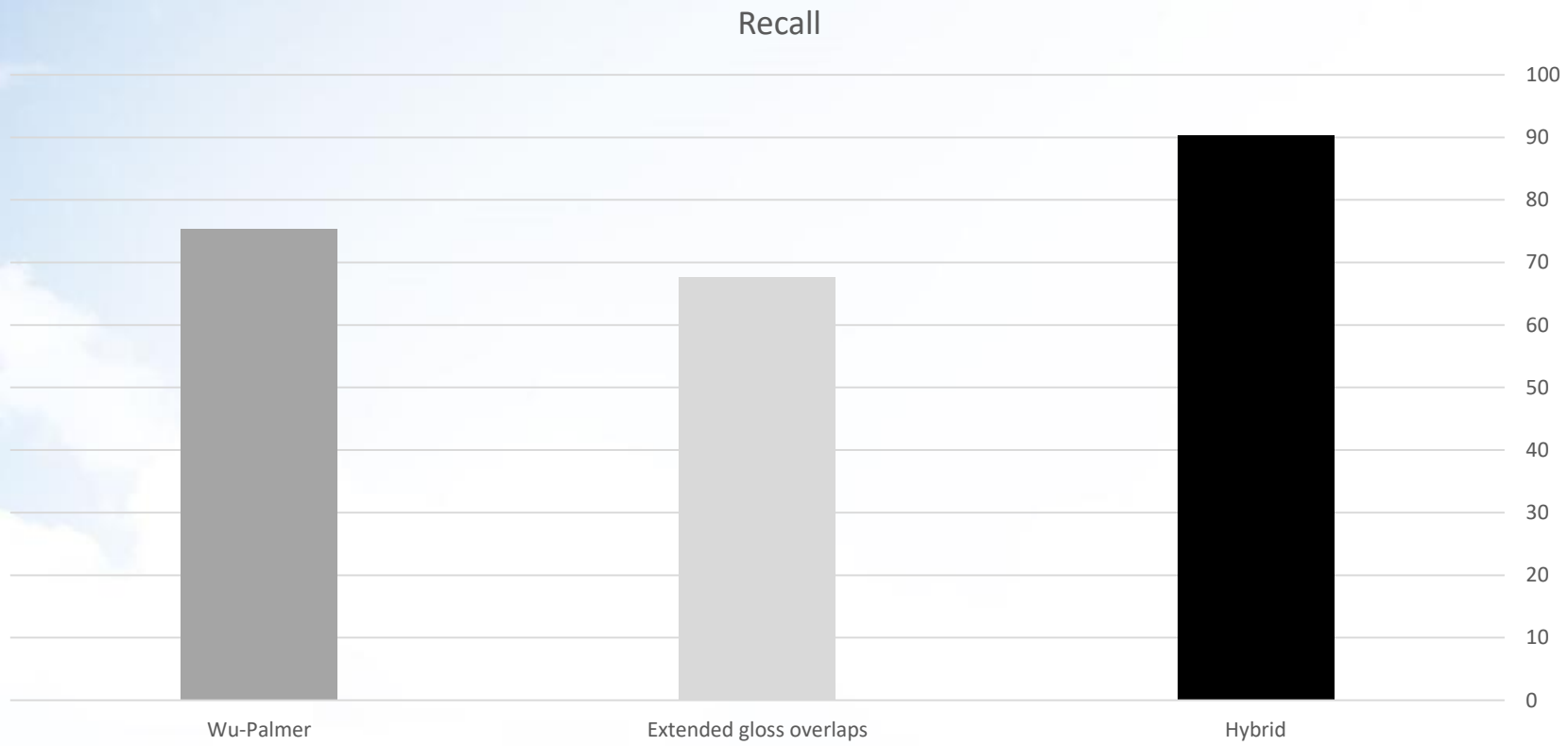
Results

- The precision of the proposed method using different semantic similarity measures (in percentage).
- Precision=What percentage of the attractions that are similar to preferences are visited?



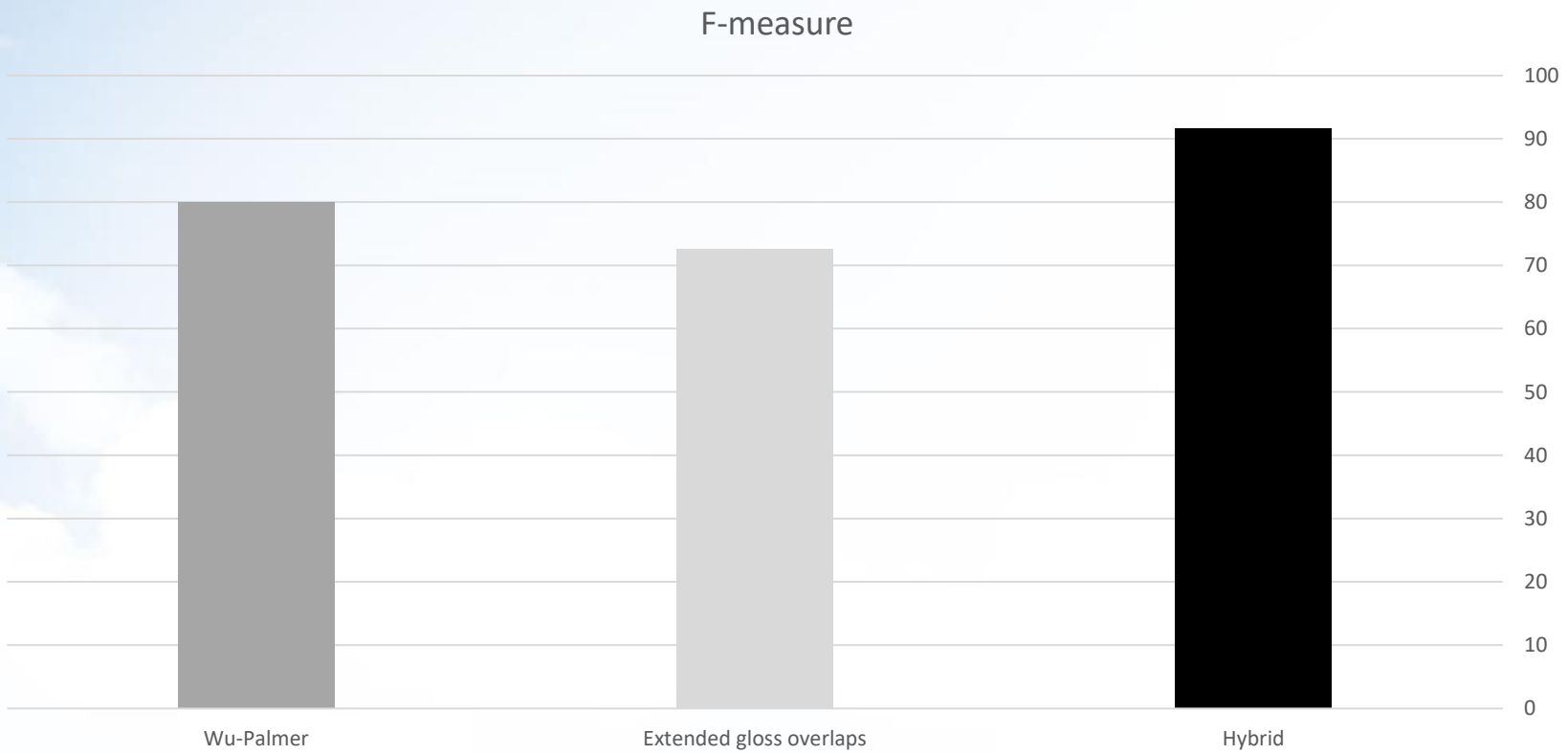
Results

- The recall of the proposed method using different semantic similarity measures (in percentage).
- Recall= What percentage of the attractions that are visited are similar to preferences?



Results

- The f-measure of the proposed method using different semantic similarity measures (in percentage).
- F-measure = $\frac{2\times precision\times recall}{precision+recall}$



Conclusion

- The proposed method has been successful in obtaining high values of evaluation parameters (precision, recall, and f-measure).

- Development of a tourism recommender system in the context of a smart city is considered as the main future work of this research.

Thank You