Towards Personalized Context-Aware Recommendation Agents in Mobile Social Networks

Santipong Thaiprayoon

Faculty of Mathematics and Computer Science
FernUniversität in Hagen, Germany

About Me

Work Experiences:

- 2005 - 2007 Master of science in information technology at KMUTNB, Thailand
- 2007 - 2020 Research assistant at NECTEC, Thailand
- 2020 Research assistant and PhD student at FernUniversität in Hagen, Germany

Research Interests:

- Natural language processing
- Information retrieval
- Machine learning
- Distributed system

Research Topics:

- Context-aware and hybrid recommendations
- Text clustering
- Keyword extraction

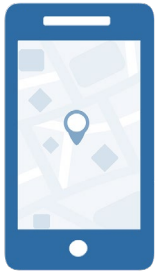
Introduction

- Context-aware recommendation agents on mobile devices are gaining popularity in the era of rapid data and technological growth.
- The agent assists in performing tasks for an individual user by discovering interesting information and facilitating the process of decision-making.

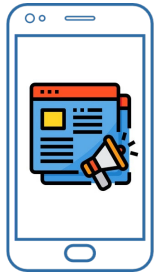
Personalized Context-Aware Recommendation Agent

- The goal is to automatically provide effective, personalized, and context-dependent recommendations based on the contextual situation of a user.
- It is designed as an autonomous intelligent agent on a mobile device to understand contextual data and adapt suggestions accurately.
- The proposed agent uses three different sources including, implicit data, explicit data, and textual data to highly improve the performance of the recommendation agent.
- The proposed agent has three challenging components.

Friend Recommendation



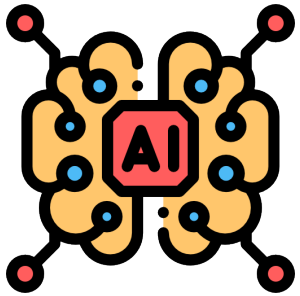
POI Recommendation



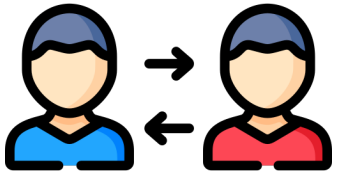
Advertisement Recommendation

Web 3.0

- The proposed agent incorporates Web 3.0 concepts to make it more personalized, context-aware, and intelligent in a decentralized manner.



Artificial Intelligence: The agent can analyze and understand contextual data like a human, using the ability of AI and NLP to provide highly personalized and relevant recommendations.



Peer-to-Peer Network: Users can communicate directly with one another without centralized authorities and other third-party intermediaries to enable the transmission of data in a secure manner.

Data Privacy and Control: Users regain full control over their data, which preserves the security and privacy of all individuals.

Data Sources

Documents

Academic Databases

Webpages

Scientific Publications

Unstructured Text

Physical Situation (Location, Weather)

Time Situation (Season, Time, Day)

Implicit Context

Profile, Bio, Interest, Questionnaire

Explicit Context

Data Acquisition

User Profile

Historical Locations

Local Server

Cloud Server

Personalized Context-Aware Recommendation

User Preference Representation

POI Representation

User-POIs Similarity Score Calculation

Advertisement Generation

Recommendation Generation

Term-based Similarity

Context-based Similarity

Keyword Extraction

Semantic Matching

Text Preprocessing
(Lowercasing, Stemming, Stopwords and Special Characters Removing)

Presentation

Friend Recommendation

POI Recommendation

Advertisement Recommendation

Bluetooth

GPS

Wi-Fi

Challenging Task#1
Friend Recommendation

Friend Recommendation

Motivation

The Corona-Warn-App is a contact-tracing app to help people by notifying them if they were recently close to infected persons with the coronavirus.

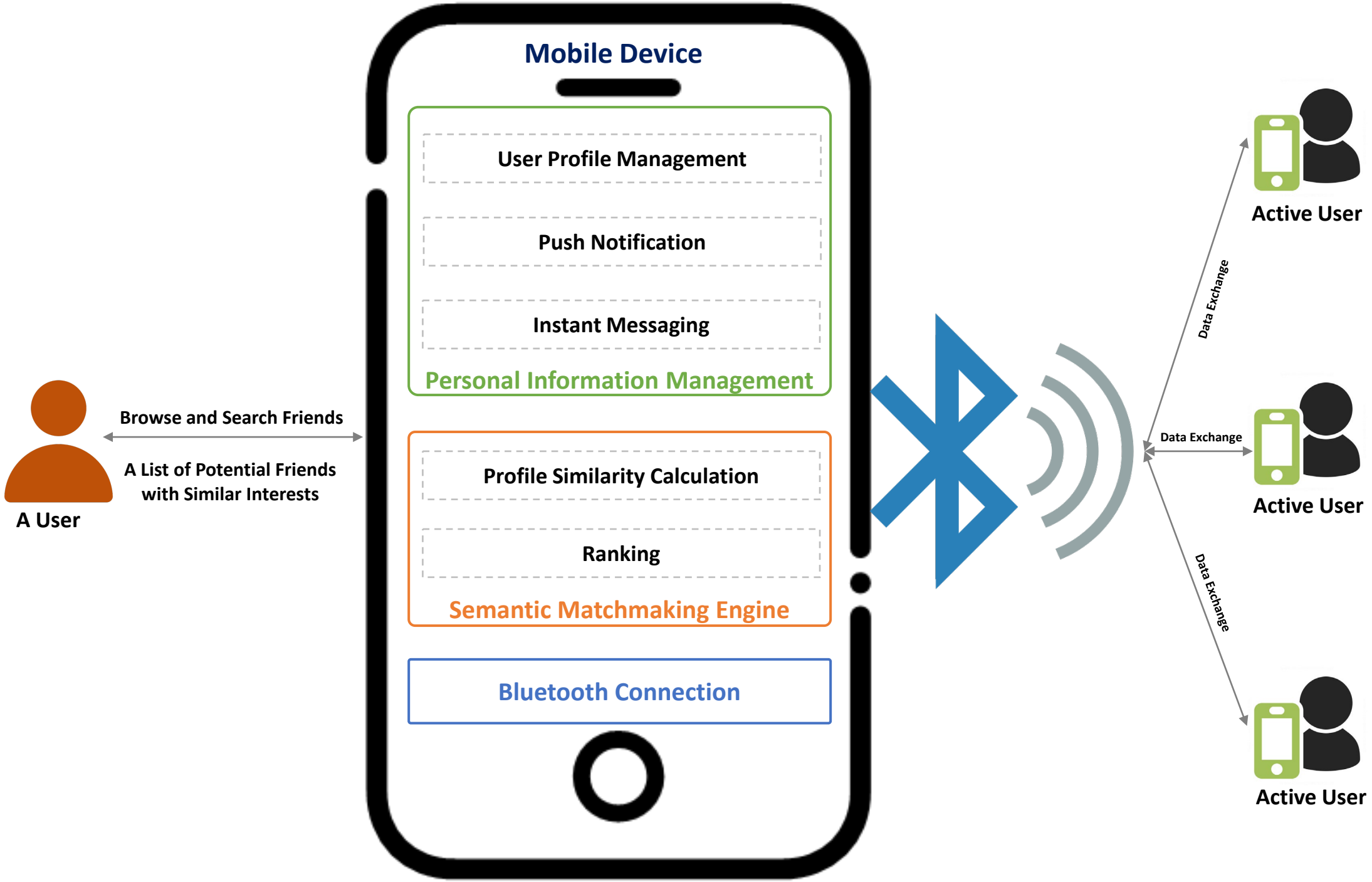
Tinder is an online dating app to meet new people everywhere by matching couples based on their profiles and locations.

Goal

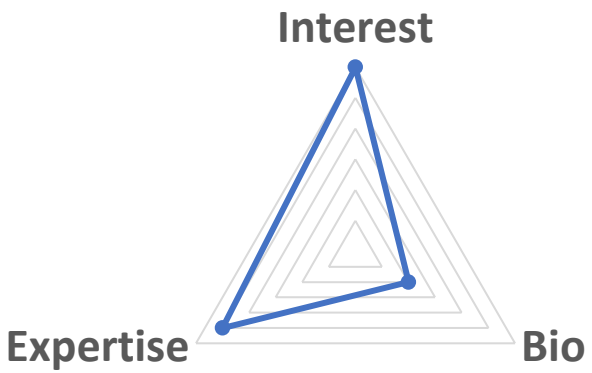
The proposed component aims to discover similar friends and provide a perfect matching within close proximity based on their preferences.

Friend Recommendation

- This component is a social networking app for meeting new friends who have common interests among nearby people using Bluetooth technology.
- The app allows users to discover new friends in their social circles.
- The personal data will be stored locally on each device preventing access and control over data.
- Users can communicate and exchange data without a central server.

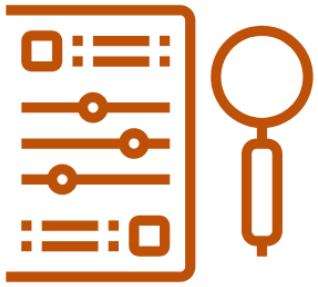

Mobile Device
User Profile Management
Push Notification
Instant Messaging
Personal Information Management
Profile Similarity Calculation
Ranking
Semantic Matchmaking Engine
Bluetooth Connection
A User
Browse and Search Friends
A List of Potential Friends with Similar Interests
Data Exchange
Data Exchange
Data Exchange
Active User
Active User
Active User

User Profiles



A User Persona

- User interests, expertise, and biographies are leveraged to calculate the similarity score.
- The user interests are manually input by tagging controlled vocabularies such as cooking, soccer, and shopping.
- The short bio is manually filled out by a user in the form of free text.
- The areas of expertise are automatically generated from documents, and academic databases via crawlers or APIs.



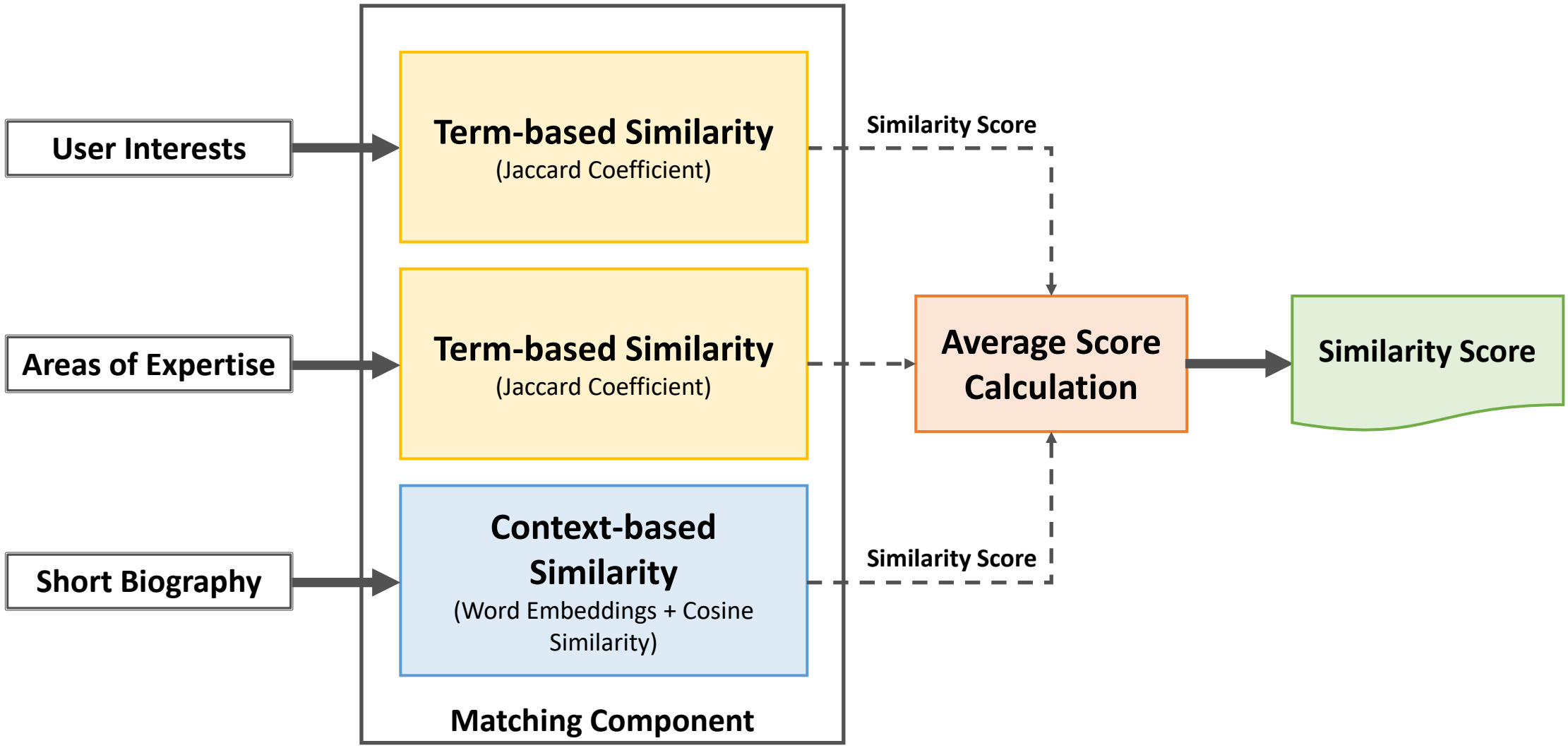
Search Filtering

- The personalities including age, gender, weight, height, and education are used for search filtering to narrow and customize search results.
- The user personalities are manually filled out by a user in the form of text fields.

Semantic Matchmaking Engine

- Semantic matching is the task of estimating semantic similarity between source and target text pieces.
- The semantic matching algorithm compares a user profile with other profiles using a text similarity technique and suggests relevant matches.
- User interests, expertise, and biographies are combined to help recommend friends to a source user.
- The text similarity technique measures the similarity score based on lexical similarity and semantic similarity covering both word level and context level.
- The algorithm can re-rank the candidates to consider probability, additional criteria, or constraints to filter some candidates and return the top ranked friends according to the similarity score to the user.

Matching Component



Term-based Similarity Based on Jaccard Coefficient

- This module aims to calculate the similarity score between two sets of words based on Jaccard coefficient.
- Jaccard coefficient is defined as the intersection and union of these two sets of words that refer to the number of common words over a total number of words.

$$\boldsymbol{J(A,B)} = \frac{|A \cap B|}{|A \cup B|}$$

Example:

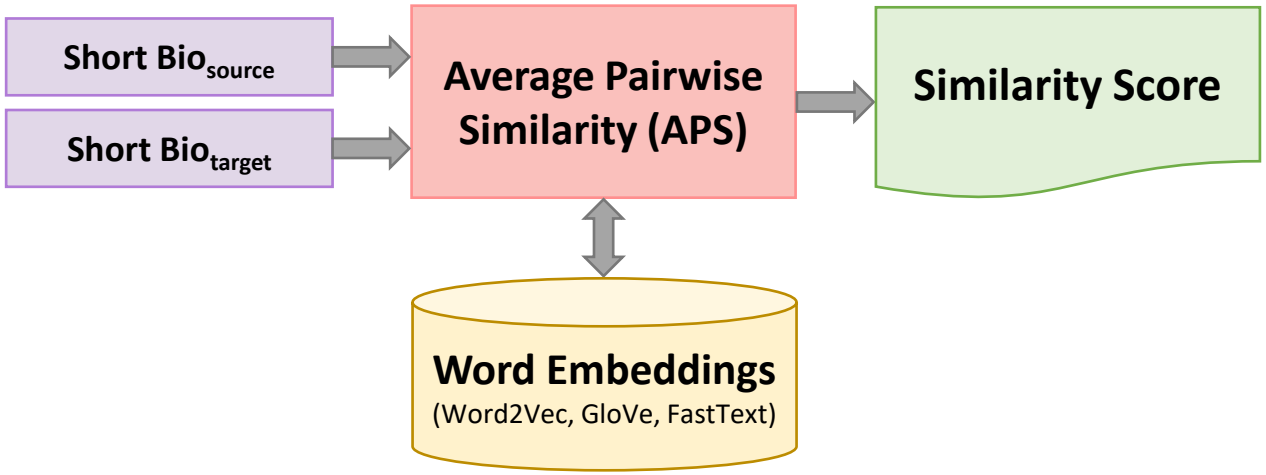
User interest$_{\textbf{source}}$ = [‘basketball’, ‘golf’, ‘vegan’, ‘travel’, ‘soccer’]
User interest$_{\textbf{target}}$ = [‘reading’, ‘writer’, ‘golf', ‘soccer’]

$$\boldsymbol{J(UserInterest_{source}, UserInteresttarget)} = \frac{\{'golf','soccer'\}}{\{'basketball','golf','vegan','travel','soccer','reading','writer'\}} = \frac{2}{7} = \mathbf{0.285}$$

Context-based Similarity Based on Word Embeddings

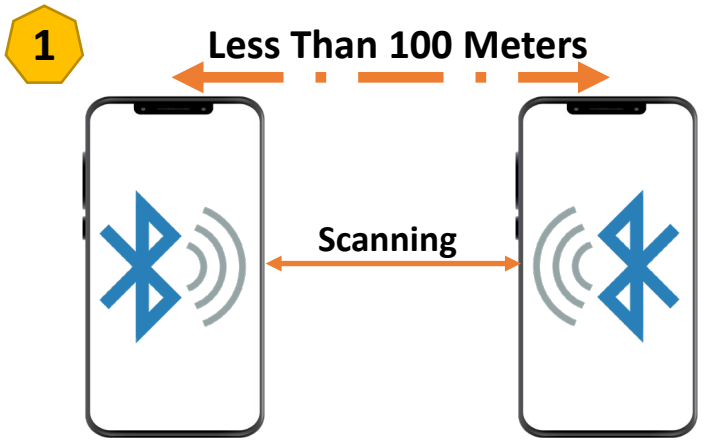
- This module aims to measure the semantic similarity of two texts on a context level.
- Average Pairwise Similarity (APS) is proposed to calculate the similarity score between two texts based on word embeddings with the assumption that similar words should have similar vectors.
- The similarity score between a word in the source text and every word in the target text is calculated as the maximum similarity score from a pre-trained word embedding model.

$$\boldsymbol{APS}(\boldsymbol{S}_{source}, \boldsymbol{S}_{target}) = \frac{\sum_{w \in S_{source}} Max\left(sim_{v \in S_{target}}(w, v)\right)}{|S_{source}|}$$

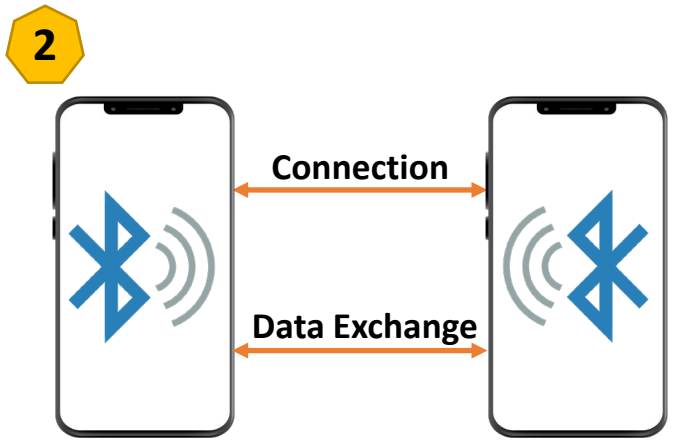


Mobile Communication Process through Bluetooth

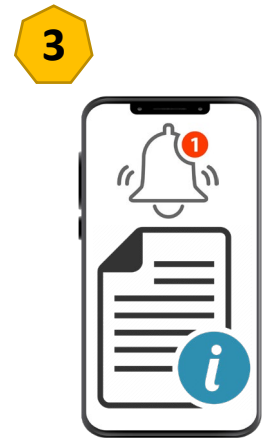
- Each Bluetooth-enabled phones periodically broadcast packets to nearby phones, allowing other phones to connect them.



When person A and person B come in close proximity



Each smartphone exchanges data via Bluetooth connection

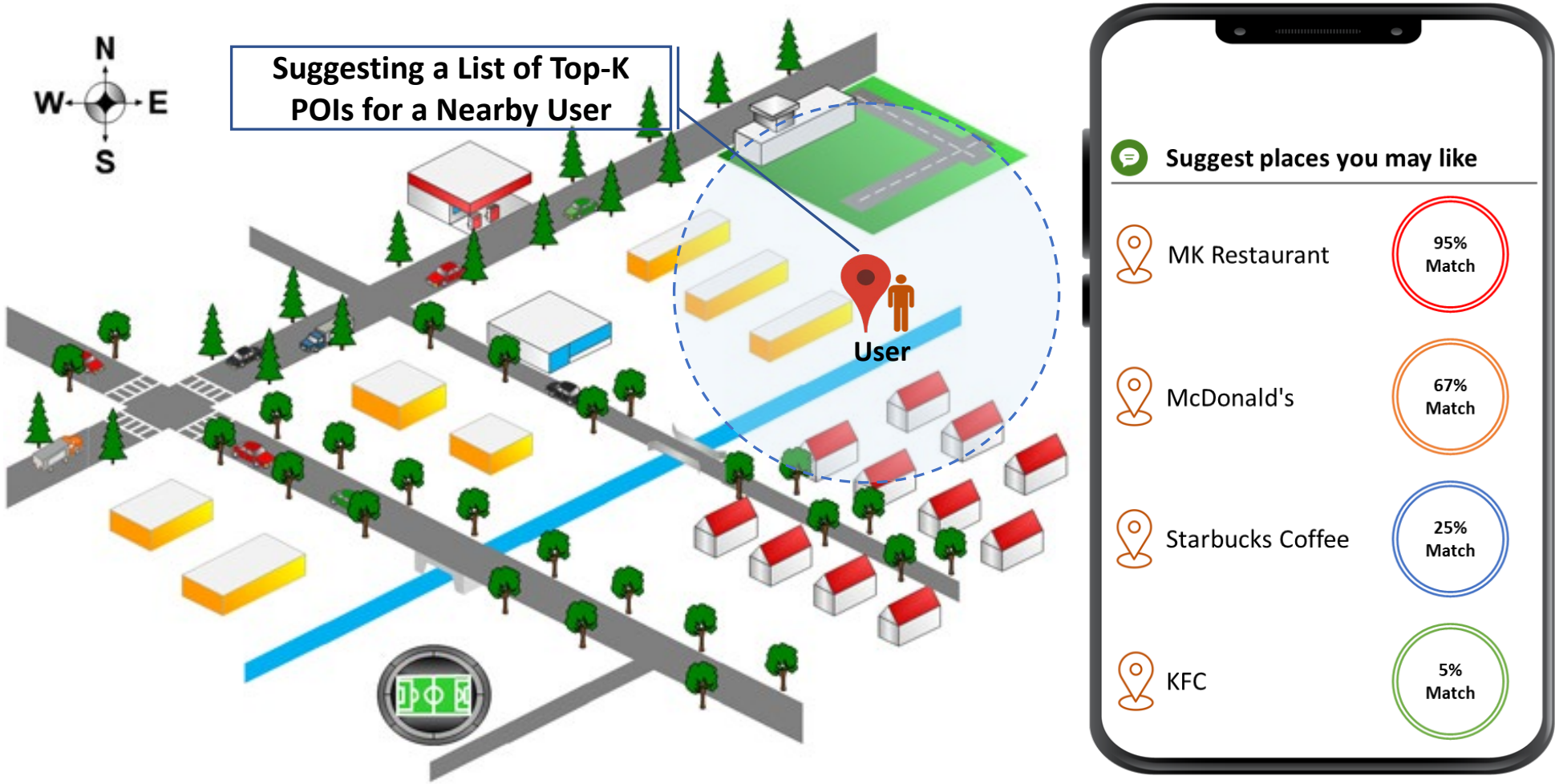


They receive notifications with information on their phones to make friends

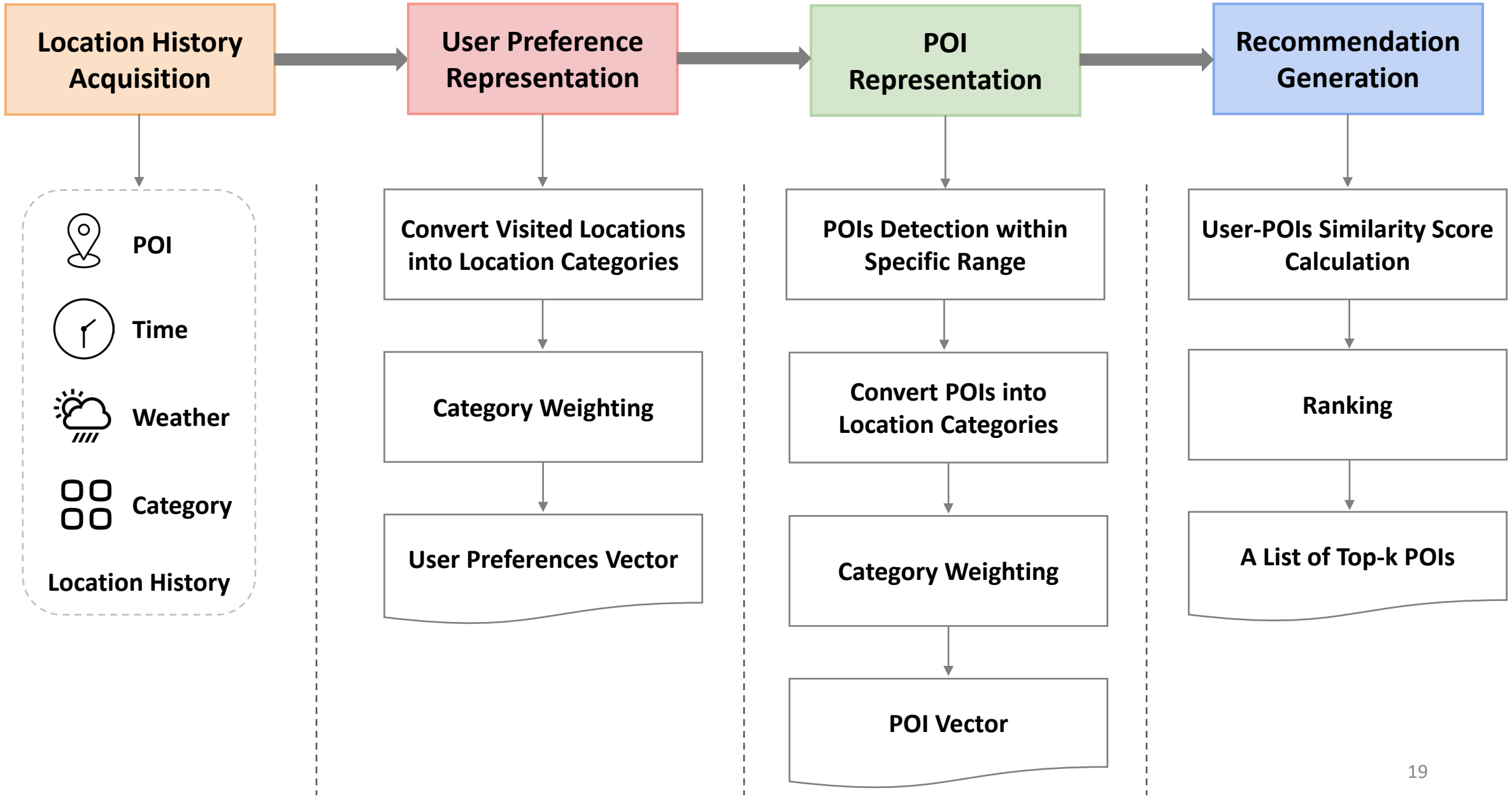
Challenging Task#2
Personalized POI Recommendation

POI Recommendation

- This component proposes a newly personalized POI recommendation app that aims to automatically recommend places of interest with in-depth information based on a user's specific interests within a predefined range from the user's current location.



Framework Overview



POI Mapping

- A POI is converted into a category sequence by location category hierarchy.
- The location category hierarchy can describe meaningful information better than POI name.
- "Dragonfly China" has a path: *Food → Asian Restaurant → Chinese Restaurant → Hainan Restaurant*.
- The path is treated as a POI profile.

Place Names	Parent Category	Subcategory 1	Subcategory 2	Subcategory 3
Dragonfly China	Food	Asian Restaurant	Chinese Restaurant	Hainan Restaurant

User Preferences Representation

- The sequential group of categories are built as a list of unique words.
- The list of unique words is represented as a numerical feature vector of user preferences.

Users	Place Names	Parent Category	Subcategory 1	Subcategory 2	Subcategory 3
User A	Narai Restaurant	Food	Asian Restaurant	Thai Restaurant	Isan Food
	Kaufland	Shop & Service	Food & Drink Shop	Supermarket	-
	Starbucks	Food	Coffee Shop	-	-
	Blue Elephant	Food	Asian Restaurant	Thai Restaurant	Spicy Salads

Categories = [‘food’, ‘asian_restaurant’, ‘thai_restaurant’, ‘isan_food’, ‘shop_service’ , ‘food_drink_shop’ , ‘supermarket’ , ‘food’ , ‘coffee_shop’, ‘food’, ‘asian_restaurant’, ‘thai_restaurant’, ‘spicy_salad’]

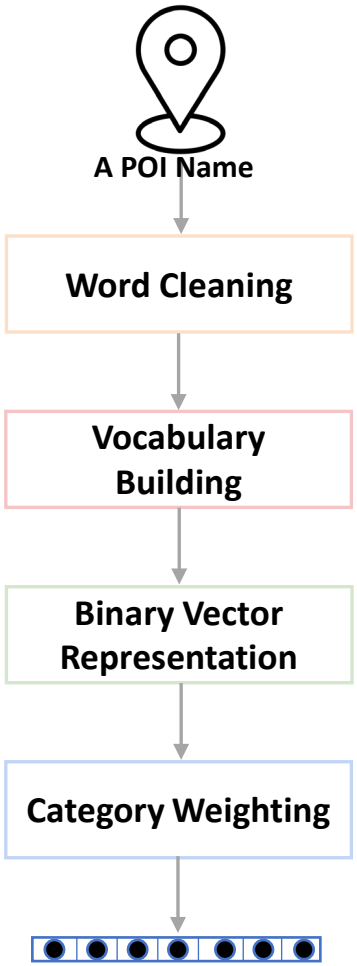
food	asian_restaurant	thai_restaurant	isan_food	shop_service	food_drink_shop	supermarket	coffee_shop	spicy_salad
1	1	1	1	1	1	1	1	1

food	asian_restaurant	thai_restaurant	isan_food	shop_service	food_drink_shop	supermarket	coffee_shop	spicy_salad
0.3	0.4	0.6	0.4	0.1	0.2	0.3	0.2	0.4

[0.3, 0.4, 0.6, 0.4, 0.1, 0.2, 0.3, 0.2, 0.4]

POI Representation

- The candidate POIs around a current location of a user are represented as a numerical feature vector defined by location category hierarchy.



Place Names	Parent Category	Subcategory 1	Subcategory 2	Subcategory 3
Dragonfly China	Food	Asian Restaurant	Chinese Restaurant	Hainan Restaurant

Categories = ['food', 'asian_restaurant', 'chinese_restaurant', 'hainan_restaurant']

food	asian_restaurant	chinese_restaurant	hainan_restaurant
1	1	1	1

food	asian_restaurant	chinese_restaurant	hainan_restaurant
0.1	0.2	0.3	0.4

[0.1, 0.2, 0.3, 0.4]

Category Weighting Model

- The category weighting model is a representation that describes the occurrence of categories within location history data of a user.
- This model assigns a weight to the significance of a particular category on the basis of the raw frequency of the categories (CF) and the weighted category level (WCL) over a whole collection of categories.
- The category frequency is the number of times that a category occurs in a category collection with the assumption that a user frequently visiting in the same place has more important.
- The top level of Foursquare category hierarchy is defined as a score of 0.1 and the bottom level has a score of 0.4 with the assumption that the bottom level can indicate the specific information of venues than the top level.

$$\boldsymbol{W(c)} = cf(c) \cdot wcl(c)$$

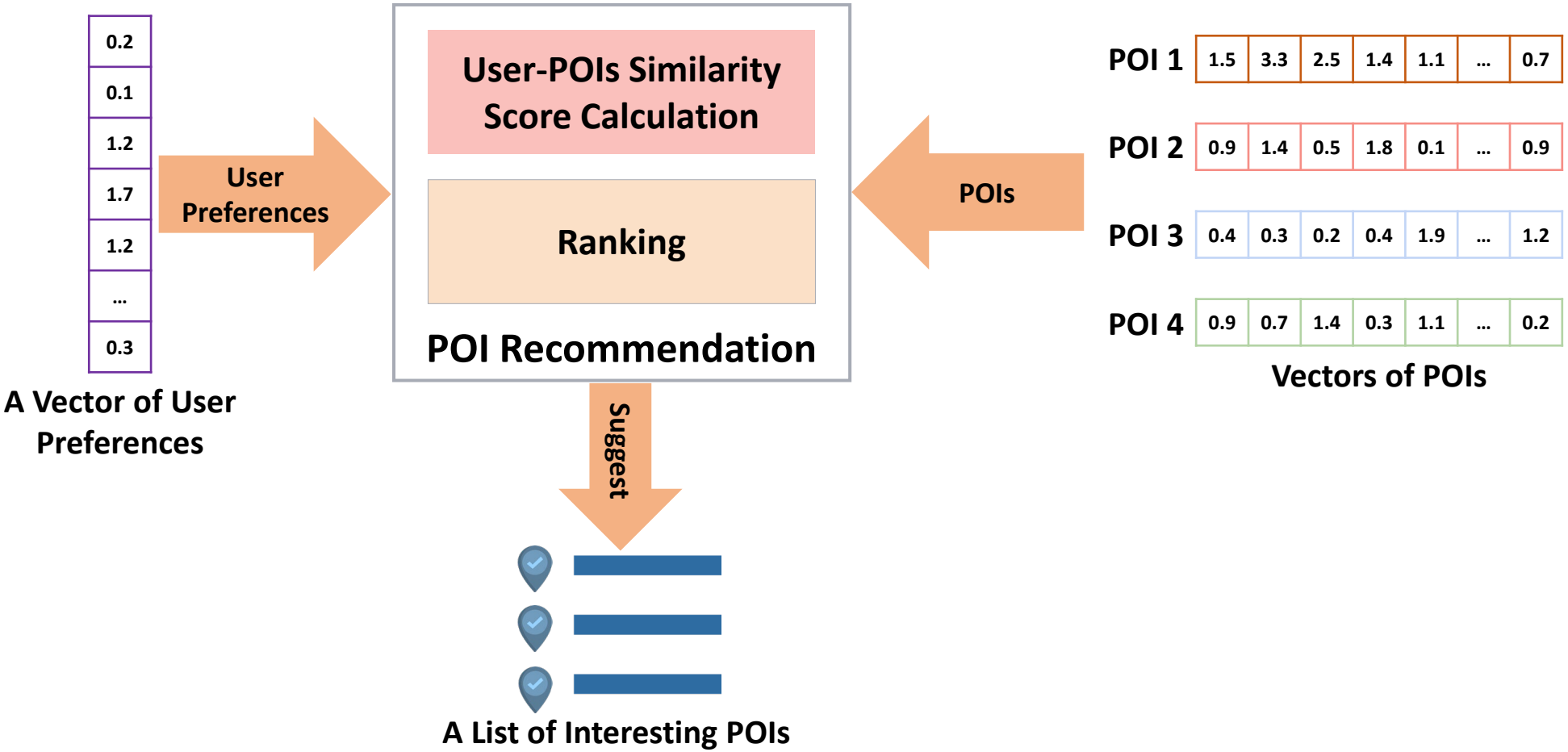
where

$$cf(c) = \frac{Number\ of\ times\ the\ category\ c\ appears\ in\ the\ category\ collection}{Total\ number\ of\ categories\ in\ the\ category\ collection}$$

$$wcl(c) = \begin{cases} 0.1, c = level\ 1 \\ 0.2, c = level\ 2 \\ 0.3, c = level\ 3 \\ 0.4, c = level\ 4 \end{cases}$$

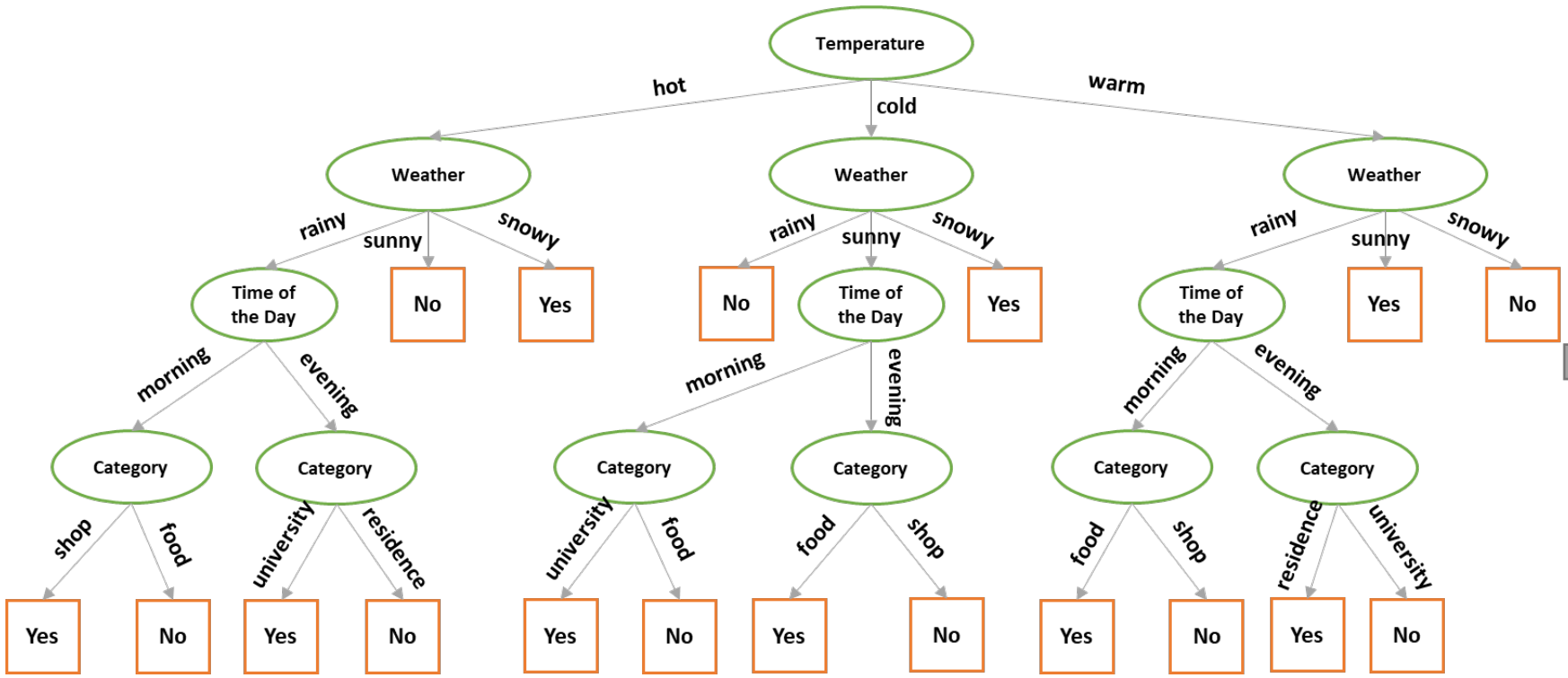
Recommendation Generation

- This component is to generate a list of POIs of interest to a target user by calculating the similarity score of two vectors.
- The vectors of a user preferences and each POI are used to determine the relevant POIs.



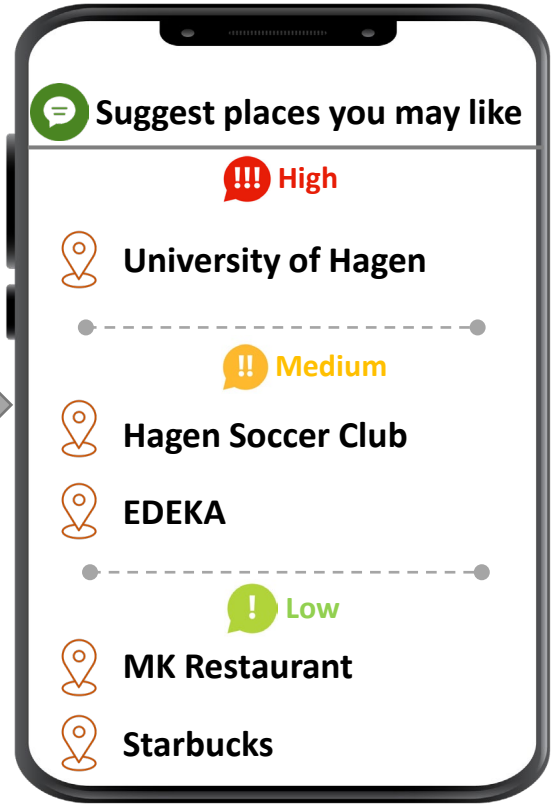
Ranking

- The intelligent recommendation should also consider other contextual factors to rank appropriate results to a user.
- The decision tree algorithm is applied to assign priority ranking on the existing results using two factors including weather conditions, and times of the day.
- The priority ranking categorizes existing results by assigning a priority of low, medium, or high.



Decision Tree Model

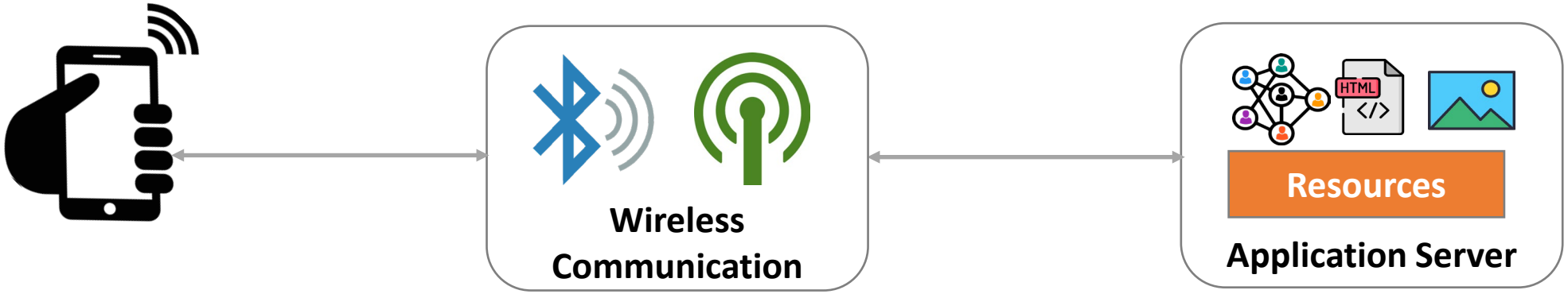
Ranking



Challenging Task#3
Advertising Recommendation

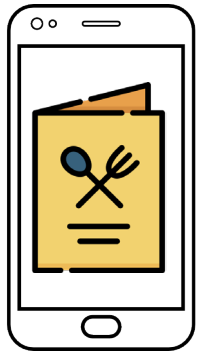
Advertisement Recommendation

- This component allows mobile users to connect to a local server via Bluetooth and Wi-Fi hotspots in different geographical regions.
- The component server automatically offers useful information or good services such as private messages, games, local recommendations, and advertising campaigns by sending push notifications to nearby users who are within the vicinity of a venue.
- The component leverages data from a mobile device of an individual user to display relevant content personalized by historical locations, weather, and interests.
- It also helps businesses provide marketing opportunities to reach target customers based on proximity at a convenient time with the right content.



Application Scenarios

- The proposed agent could be applied to smartphone users for receiving useful information based on profiles within a predefined area via geolocation technologies.

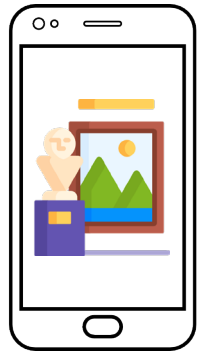


Restaurants

Attract target customers with advertising campaigns such as coupons, special promotion, and discount

Airports

In addition to flight information, showcase duty-free promotions and special vacation deals



Museums

Provide mobile tours for exhibitions and guide visitors with floor plans, events, and shop promotions

Customers who enter a shop and open an app connected to the shop's free Wi-Fi, they will receive notifications on their smartphones about promotions and menu items related to the shop.

When audiences enter an academic conference and open an app connected to a local Bluetooth, and Wi-Fi network, they will receive push notifications about available scheduled programs in the academic conference.

Future Work

- We plan to enhance the proposed agent as an intelligent personal assistant, such as Apple Siri, Amazon Alexa, and Google Assistant.

Apple Siri Amazon Alexa Google Assistant

Voice Commands: Communicating with people in complex sentences and responding with useful answers using natural language.

Intelligent Interpretation: Understanding contextual data of a user such as current location, personal data, contact data, or past frequent locations and recommending places or items to the user.

Privacy: Storing private and user-sensitive data in a mobile device to prevent unauthorized users to access resources and services.

Conclusion

- A personalized context-aware recommendation agent is proposed.
- The agent can perform tasks for an individual user based on a combination of user preferences and location awareness in the current environment.
  - Friend recommendation
  - POI recommendation
  - Advertisement recommendation
- AI techniques and NLP are used to deeply understand contextual data.
- Web 3.0 is enabled through decentralization and privacy.
- Personal data are stored in a local database to preserve data privacy.
- GPS, Wi-Fi, and Bluetooth technologies are adopted to determine the location of a user and provide location-based information.


Thank you