CAT 400: UNDERGRADUATE MAJOR PROJECT (Final demo)

EXPERT SEARCH II: THE SEARCH

EXAMINERS   : Professor Dr. Mandava Rajeswari

                          Associate Professor Dr. Cheah Yu-N
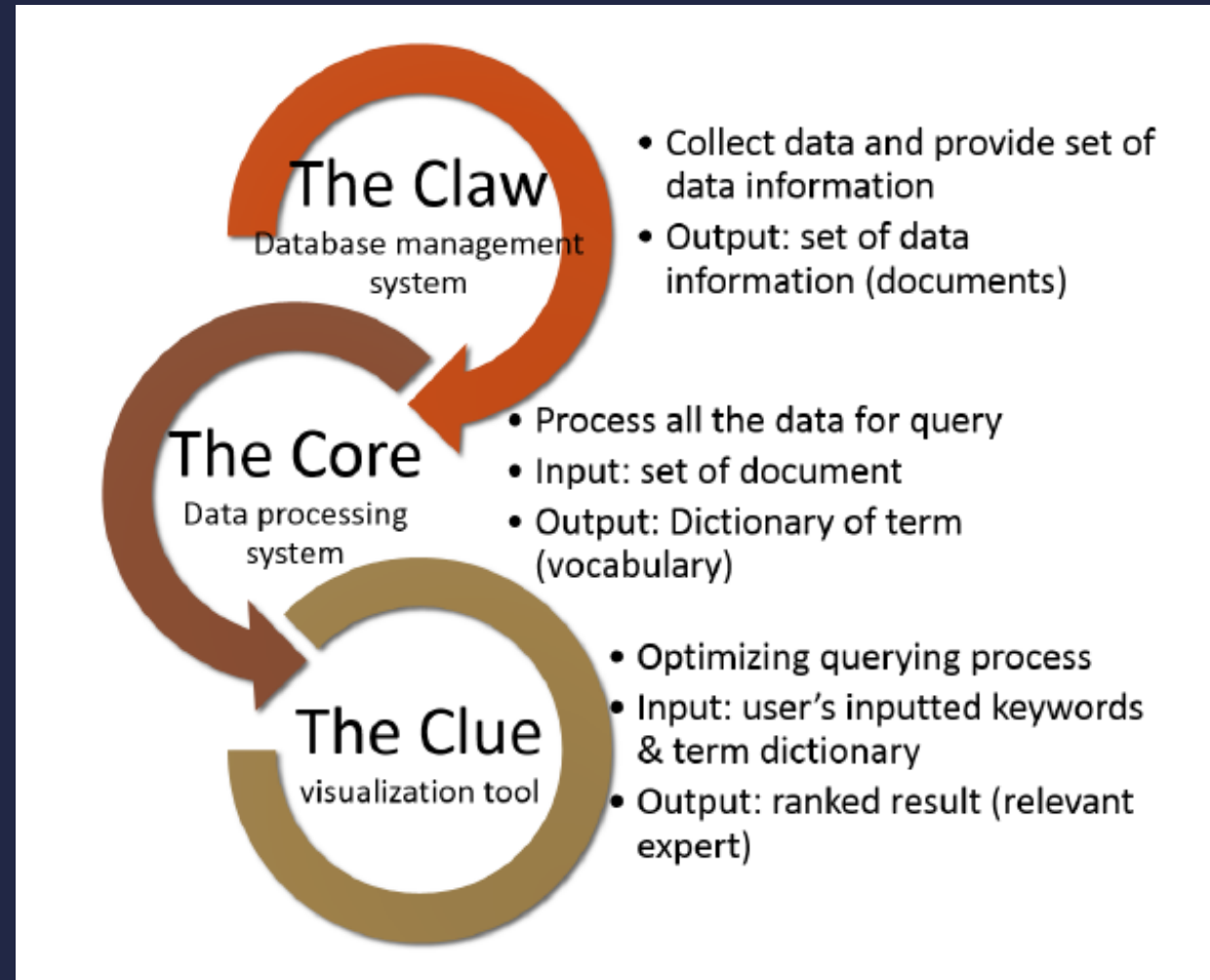
SUPERVISOR   : Dr. Gan Keng Hoon

PRESENTED BY   : Teh Chek Wei (115140)

WHAT IS EXPERT SEARCH

- Web-based search engine system designed to search for experts in academic or research organization settings.

- Expertise Scope: Publications of experts.

- Domain Scope: USM School of Computer Sciences academic staffs

- Data Scope: DBLP Indexed Database

STRUCTURE OF EXPERT SEARCH 1 (FYP 2014/15)

The Claw — Database management system
- Collect data and provide set of data information
- Output: set of data information (documents)

The Core — Data processing system
- Process all the data for query
- Input: set of document
- Output: Dictionary of term (vocabulary)

The Clue — visualization tool
- Optimizing querying process
- Input: user's inputted keywords & term dictionary
- Output: ranked result (relevant expert)

OBJECTIVE OF EXPERT SEARCH 2

- To improve the result filtering in the current system

- Help end-users to navigate results in a faster way

- Increase the coverage of Experts Data

- Help users to get more results they needed

EXPERT SEARCH 2
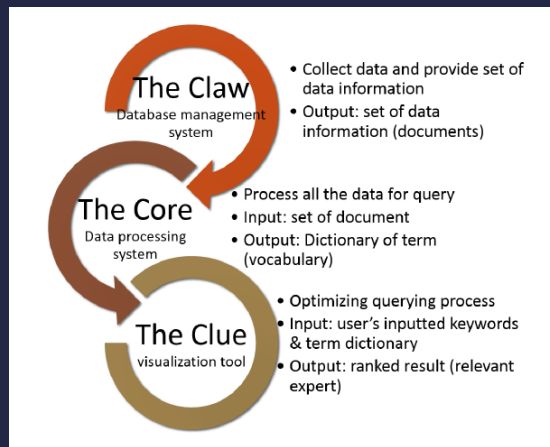
THE SEARCH (TEH CHEK WEI)
- ADD **FACET** FILTERING FUNCTION IN SEARCH FUNCTION
- DETECT ALL THE **FACET** VALUE
- PROCESS AND INCREASE THE RELIABILITY OF **FACET** VALUE
- OPTIMISE FACET VALUE BY FREQUENCY
- FLEXIBLE TOOL FOR OTHER DOMAIN

THE SOURCE (HEW HUANG)
- DETECT THE LECTURER'S ARTICLE IN **GOOGLE SCHOLAR**
- GET THE INFORMATION OF THE ARTICLE

THE SEARCH

- Facet Value Extractor
- Facet Value Optimizer
- Facet Editor
- Facet Value Filtering

ADMIN

END-USER

LET US START WITH EXPERTISE SEARCH

- EXPERTISE SEARCH IS A PROCESS OF FINDING INFORMATION RELATED TO EXPERTS IN THEIR AREAS OF EXPERTISE

- ORGANIZATION LIKE UNIVERSITIES, RESEARCH CENTERS, COMPANIES

EXPERTISE SEARCH: MOTIVATION

- HOW DO WE IMPROVE THE DISCOVERY OF EXPERTISE?
  - Answer: FACETED SEARCH
- WHAT IS FACETED SEARCH?
  - FACETED SEARCH IS ALLOWING USERS TO EXPLORE A COLLECTION OF INFORMATION BY APPLYING MULTIPLE FILTERS
    - YEAR
    - VENUE
    - COAUTHOR
    - …

PROBLEM

- THE CURRENT FACETS ARE GENERAL FILTER

- A FACET AND ITS VALUE IS NORMALLY GIVEN

- FOR EXAMPLE:
  - "YEAR" FACET AND ITS VALUE "2009"
  - "AUTHOR" FACET AND ITS VALUE "SYAHEERAH L. LUTFI" , "JUAN MANUEL MONTERO" AND OTHERS

```xml
▼<dblp>
  ▼<inproceedings key="conf/biostec/LutfiMBLG09" mdate="2014-12-12">
      <author>Syaheerah L. Lutfi</author>
      <author>Juan Manuel Montero</author>
      <author>Roberto Barra-Chicote</author>
      <author>Juan Manuel Lucas-Cuesta</author>
      <author>Ascensión Gallardo-Antolín</author>
    ▼<title>
        Expressive Speech Identifications based on Hidden Markov Model.
      </title>
      <pages>488-494</pages>
      <year>2009</year>
      <booktitle>HEALTHINF</booktitle>
      <crossref>conf/biostec/2009hi</crossref>
      <url>db/conf/biostec/healthinf2009.html#LutfiMBLG09</url>
  </inproceedings>
</dblp>
```

PREVIOUS SYSTEM (ES1) VS CURRENT SYSTEM (ES2)

HOW DO WE GET THIS???
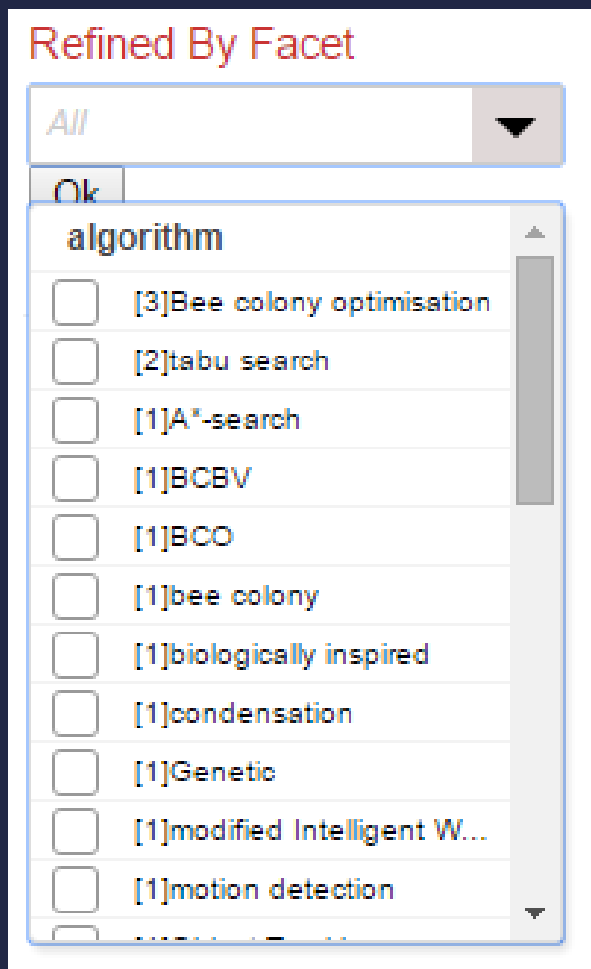
Refined By Facet

All ▼

Ok

algorithm

- ☐ [3]Bee colony optimisation
- ☐ [2]tabu search
- ☐ [1]A*-search
- ☐ [1]BCBV
- ☐ [1]BCO
- ☐ [1]bee colony
- ☐ [1]biologically inspired
- ☐ [1]condensation
- ☐ [1]Genetic
- ☐ [1]modified Intelligent W...
- ☐ [1]motion detection

FACET VALUE EXTRACTOR (ADMIN)

- A TOOL THAT HELP ADMIN TO EXTRACT OUT ALL FACET VALUE

System demo

METHOD TO EXTRACT THE FACET'S VALUE

1)GET TITLE AND ABSTRACT OF ARTICLE

Design of an Intelligent Route Planning System Using an Enhanced A*-search Algorithm.

Poh Lee Wong , Mohd Azam Osman , Maziani Sabudin

Abstract

Traffic congestion has become a major problem in many countries. One of the main causes is the failure to manage a journey. Vehicles tend to choose the shortest paths which ended up congesting a certain area. If a journey is well managed, using alternative paths may lead to the same destination in a shorter period of time. Therefore, with the importance of sustainable environment, intelligent route planning system has become a popular research agenda among researchers and industries. In this paper, an enhanced algorithm based on A*-search algorithm is proposed. The enhanced algorithm is used to study the traffic congestion level as well as to suggest the best route to a destination. The algorithm takes the traffic congestion level into consideration, which makes choosing the best route possible in minimal travel time.

METHOD TO EXTRACT THE FACET'S VALUE

2)SPLIT TEXT CONTENTS INTO SENTENCE LEVEL

Sentence:

1. Design of an Intelligent Route Planning System Using an Enhanced A*-search Algorithm.
2. Traffic congestion has become a major problem in many countries.
3. One of the main causes is the failure to manage a journey.
4. Vehicles tend to choose the shortest paths which ended up congesting a certain area.
5. If a journey is well managed, using alternative paths may lead to the same destination in a shorter period of time.
6. Therefore, with the importance of sustainable environment, intelligent route planning system has become a popular research agenda among researchers and industries.
7. In this paper, an enhanced algorithm based on A*-search algorithm is proposed.
8. The enhanced algorithm is used to study the traffic congestion level as well as to suggest the best route to a destination.
9. The algorithm takes the traffic congestion level into consideration, which makes choosing the best route possible in minimal travel time.
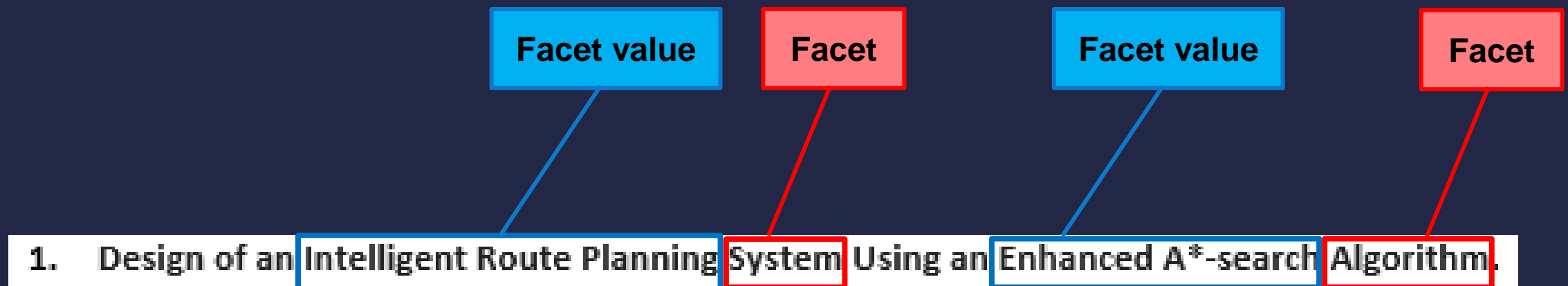
METHOD TO EXTRACT THE FACET'S VALUE

3)DETECT FACETS (e.g. Algorithm, Model)

4)DETERMINE FACET'S VALUE WITH ASSISTANCE OF THE STOPWORD LIST WHICH GET FROM ONLINE SOURCE

5)PROCESS FACET VALUE

6)STORE VALUE IN DB TABLE

Facet value **Facet** **Facet value** **Facet**

1.   Design of an Intelligent Route Planning System Using an Enhanced A*-search Algorithm.

PROCESS THE FACET'S VALUE

- REMOVE OR HANDLE THE SYMBOL APPROPRIATELY (e.g. **? : ; . ! { }**)

- HANDLE AND SPLIT THE FACET VALUE WITH "AND" & COMMA CONDITION

- REMOVE DUPLICATE FACET VALUE ON THE SAME FACET AND ARTICLE

- CAPITAL DETECTION

- STANDARDIZATION OF WORD USAGE (BRITISH WORD IS USED AS STD)

pure MPI, pure OpenMP and hybrid MPI-OpenMP =>
1=pure MPI, 2=pure OpenMP, 3 = hybrid MPI-OpenMP

BRITISH WORD	AMERICAN WORD
accessorise	accessorize
analyse	analyze
behaviour	behavior
categorise	categorize
fibre	fiber
tonne	ton
standardisation	standardization

FACET VALUE OPTIMIZER (ADMIN)

- Focus on the frequency

- Detect facet value without facet side by side

- Eliminate the facet value which is occurs too frequent in all publication

System demo

METHOD TO **DETECT** FACET VALUE WITHOUT FACET

- LEARN THE PATTERN FROM THE FACET VALUE PAIRS WHICH ARE GENERATED FROM THE FACET VALUE EXTRACTOR

- FROM THE PATTERN, WE CAN USE THE PATTERN TO DETECT THE FACET VALUE WITHOUT FACET SIDE BY SIDE IN THE OTHER PUBLICATION

secondly a sophisticated computational technique that mimic survival and natural processing which is called as Artificial Bee Colony as feature selection

PROCESS

- Calculate the frequency
- Display to user
- User set number limit to insert the facet value and eliminate facet value
- New detected facet value will assigned to highest facet count on the particular facet value which extracted from facet value extractor

Problem

How to resolve the problem on define the facet on other domain?

How to make this system flexible to OTHER DOMAIN?

FACET EDITOR(ADMIN)

- Insert facet and facet stem

- Delete facet stem

- The facet and facet stem list is prepared by the expert in specify domain

Network, networks, networking

System demo

FACET VALUE FILTERING (END USER)

- Deployed facet value extracted from the admin side to front end

- User able to refine publication by facet value

- Facet value which presented is in facet group form

System demo

CONCLUSION

- DYNAMIC FACETED SEARCH
- RUN ON REAL DATASET
- FACET VALUE OPTIMISER(FREQUENCY)
- TOOL FOR EXTEND TO OTHER DOMAIN

Q&A

THANK YOU