**컴퓨터과학종합설계 10주차 보고서**

|  |  |
| --- | --- |
| **팀명** | 4조 |
| **팀장** | 2015920021 박인규 |
| **팀원** | 2017920011 김은해 / 2017920036 양다은 / 2018920059 허정우 |

|  |  |
| --- | --- |
| **회의 일시** | 2021.11.7 일요일 14:00 – 15:00 |
| **회의 장소** | Google Meet |
| **회의 결과** | |
| * 클러스터링 방식 조사 및 논의 * 공인 IP를 할당하여 서버간 연동 준비 * 담당 파트 개발 진행 | |
| **회의 내용** | |
| - 클러스터링 알고리즘 논의   1. DBSCAN(밀도기반 클러스터링)    * 영역정보를 격자형 점 정보로 변환한 뒤에 일반적인 point 클러스터링 알고리즘을 적용해서 군집을 만들고 그걸 다시 영역 정보로 적용    * 작은 데이터셋 뿐만 아니라, 대용량 공간데이터에 대해서도 효율적으로 군집화 연산이 가능    * 알고리즘 과정    1. 임의의 점 선택 (어떤 값이든 동일한 결과 출력)    2. 해당 에서 만큼 떨어진 공간이 개수를 만족시키는지 확인      * + 1. 만족시킨다면 가 다시 만큼 떨어진 공간에 개수를 만족시키는지 확인     2. 만족시키지 못한다면 다시 임의의 점 선택   1. 를 확장시켜 나가면서 를 만족시키지 못하는 경계 포인트(border point)가 얻어질 때까지 진행   + https://rfriend.tistory.com/587  1. CLARANS    * PAM과 CLARA을 결합한 알고리즘    * 대규모의 데이터베이스에서의 분할 방법인데, DB전체 분할에 부담이 있을 경우 원본DB에서 표본을 추출하고 추출한 표본으로부터 가장 좋은 군집 출력    * https://glorymind.tistory.com/entry/kmedoids%EC%A4%91%EC%95%99%EA%B0%9D%EC%B2%B4-%EB%B0%A9%EB%B2%95    * https://www.koreascience.or.kr/article/JAKO200323421079395.pdf 2. 격자기반 클러스터링    * 처리속도가 빠르다는 장점    * 하지만, 대량의 공간 데이터 집합을 클러스터링 하는 것은 데이터 처리 비용이 급격하게 증가하기 때문에 클러스터링 처리 결과에 큰 영향 미침    * 김호석, Song Gao, 김경배, Ying Xia, 배해영. (2006). 대용량 공간데이터베이스를 위한 확장된 밀도-격자 기반의 공간 클러스터링 알고리즘. 정보처리학회논문지D, 13(6), 633-640. 3. Fuzzy clustering    * ‘군집에 속할 신뢰성이 어느정도인가’에 대해 산출    * 참, 거짓으로 명확히 구분하는 대신, A군집 80%, B군집 30%와 같이 유연하게 처리    * 잡음과 잘못된 제보에 대한 처리에 강인함   - 네이버 클라우드 플랫폼 공인 IP 할당하여 개인 IP 배분   * DB서버에서 접속충돌 발생 -> 오류 검토중   **-** 각 서버 진행상황 보고 및 검토 진행  - 추가적 구현 페이지(논의중) : 관리자 페이지 구축 | |
| **향후 계획** | |
| - 구현할 클러스터링 알고리즘 최종결정 및 적용  - 각 파트 구현 완료 및 보고  - 시스템, API, DB 서버 연동(연동과정에서 오류와 테스트 수시 진행) | |