캡스톤 18조 회의록

프로젝트명	O24Sec	팀명	멜러리를 찾아서
회의 시간	4월 23일 16:00 - 20:00		
장소	국민대학교 미래관 7층 K랩		
참석자	김민송, 장우혁	서기	장우혁
안건	 졸프 실험 범위 재정의 데이터 관리 및 유사도 검사 멘토링 및 추후 계획		
	조교 사람 많이 제작이		

졸프 실험 범위 재정의

윈스, 지도교수님과의 미팅 결과를 토대로 3차 목표인 유사도 검사 및 클러스터링에 대한 실험 범위 재정의

- 회의(멘토링 보고서에 첨부) 결론: 1차, 2차 구현 목표에 만족하고 추가적으로 비슷한 선례인 KISTI실험의 재구성을 목표
- 시나리오 재구성을 위해 KISTI 실험 분석 필요
- KISTI 실험 당시, 보안 관제 요원들이 판단한 100% 정확한 패킷 종류 1개를 기준으로 나머지 패킷에 대한 유사도 분류 후, 진행 -> 보안 관제 요원이 없는 우리 상황에 맞춘 해결법이 필요
- 문장 기준을 패킷으로 설정하여 자연어 처리 기법 등을 사용 예정

데이터 관리 및 유사도 검사

최대 2천만 건의 패킷 이벤트 데이터를 유사도 검사하기 위해선 데이터 정제 및 분류 그리고 추가적인 처리 기법 적용이 필요함

- 2천만 건의 패킷 데이터에서 특정 포트(80) 서버의 통신 이벤트로만 제한
- 회의 내용
- 기간을 특정 개월(5 ~ 9월)로 제한
- 내부 서버 이벤트(라벨링 1번)으로 제한
- Top-k기법으로 feature의 개수를 줄이고, 패킷 데이터를 feature크기가 아닌, index count방식으로 저장

유사도 검사에서는 전체 문장에 대한 유사도 검사를 진행 예정, 해결해야 할 문제

- 유사도 기준을 어떻게 잡을 것인지
- 정탐과 오탐 사이의 기준에 차이를 둘 것인지 이상 2가지 기준은 실험 결과 분석 후 진행

멘토링 및 추후 계획

바뀐 3차 목표에 대한 Ceeya 멘토님과의 멘토링 진행 예정(유사도 검사 기법 및 데이터 선정 방식 등): 4월 30일 오후 2시 예정

현재 진행중인 로드맵(아래 '프로젝트 진행사항' 참고)에서 객체별 클러스터링 기법 실험 부분에 해당하는 내용이 변경되었지만, 정해진 로드맵에서 바뀔건 없으므로 일정 스케쥴은 그대로 진행,

Github 코드 리팩토링, 자료 추가 작성 등을 병행하면서 진행

- 추가 자료: 최종 발표때 사용할 PPT, 최종 보고서, 포스터, 짧은 소개 동영상 등
- 링크 지원금 사용 적극 추진해야 함

