**Best of Best 11**

**Gartner 2022**

**김영옥 멘토님 과제 #1 번**

**Best of Best 11기**

**보안제품 개발 희망 정경재**

**제출날짜:2022-07-13**

**목차**

1. **Gartner 2022 Privacy Enhancing Computation을 고른 이유**
2. **Privacy에 대해여**
3. **PEC에 대해서**
4. **PEC가 필수로 자리잡을 이유**
5. **PEC의 기술들의 장-단점**
6. **PEC의 활용방안**
7. **공부하며 느꼈던 점들**

저는 이번 Gartner에서 발표한 전략기술 트랜드 2022에서 제가 가장 관심이 있는 분야는 Engineering Trust(신뢰 구축) 부문의 Privacy-Enhancing Computation PEC이라고 불리는 프라이버시 강화 컴퓨테이션에 대해 가장 관심이 있습니다.

제가 다양한 12가지의 전략기술 중에서 Privacy-Enhancing Computation에 가장 관심을 가지게 된 이유는 저는 보안의 기본적인 토대라고 생각하는 인간의 기본적인 권리 중에서 가장 기본이라 생각하는 것이 “프라이버시”라고 생각하기 때문입니다.

저의 이러한 생각은 제가 프라이버시의 발전과정을 배우고, 이러한 침해를 몸으로 느끼며 이러한 생각을 가지게 되었습니다.

프라이버시의 시작은 미국의 인기 변호사의 권리를 보호받기 위함으로부터 시작되었습니다. 1888년 당시 미국의 인기 변호사였던 워렌(Warren)과 그의 부인은 황색언론을 통해 그들의 사생활이 보도되면서 그들의 사생활이 언론에 보도되지 않을 권리인 ‘혼자 있을 권리(right to be let alone)’ 주장하게 되었고 법정은 이러한 권리에 대해 보호받아야 함이 마땅하다며 워렌의 손을 들어주게 되면서 시작되었습니다.

이후 1888년도의 ‘혼자 있을 권리(right to be let alone)’를 1890년도에 워렌(Warren)과 브랜다이스(Brandeis)가 privacy로 구체화하였습니다. 이렇게 privacy의 개념이 대두되고 난 이후 세계는 정보, 통신 기술의 발전을 이루게 되었습니다.

이러한 기술들의 발전에 의해 기존의 privacy만으론 스스로 공개한 개인정보나 사생활에 대해 통제할 수 없게 되면서 개인정보나 사생활에 대해 통제가 필요가 대두되었고 기존의 권리에 ‘개인정보 자기 결정권’인 적극적인 성격의 권리로 확장되게 되었습니다.

이에 따라 사람들은 자신에 관한 정보를 자율적으로 결정하고 관리할 수 있는 현대의 privacy의 개념으로 발전하게 됩니다.

이렇게 적극적인 성격의 권리로 확장되며 사람들은 다양한 매체를 통해 이전까지 자신의 권리들이 얼마나 보호받지 못했는지 인지하게 되었고, 추가로 이렇게 노출된 개인들의 정보로 큰 피해를 받게 되면서 점차 사람들은 점점 자신의 privacy가 왜 보호받아야 하는지에 대해 인지하게 되면서 자신의 privacy가 보호받아야 함을 적극적으로 주장하기 시작했습니다.

이렇게 privacy가 대두됨에 따라 개인들은 기업에 자신의 Privacy를 지킬 수 있는 까다로운 정책을 요구를 제시하게 되었습니다. 기업들은 이러한 개인의 요구와 법적인 제재에 부합하기 위해 점점 진화된 privacy 보호 정책 등 다양한 방법을 고민하고 privacy를 높일 수 있는 방법을 제공하기 시작했습니다.

따라서 이렇게 privacy에 대해 학습하고 실제로 개인적으로 privacy를 침해 받아 문제가 발생했던 경험과 다수의 기업이 사용자들의 privacy 문제로 인해 곤란을 겪는 것을 보면서 저는 privacy가 앞으로도 가장 중요한 보안이슈라고 생각하게 되어 이번 Gartner의 12가지 주제에서 Privacy-Enhancing Computation을 선택하게 되었습니다.

Privacy-Enhancing Computation(프라이버시 강화 컴퓨테이션)은 아직까지는 명확한 정의는 없지만, 소프트웨어와 하드웨어 기반에서 개인정보를 훼손하지 않고, 민감한 정보와 데이터를 안전하게 공유하고 암호화 및 분석할 수 있는 기술을 의미한다고 말합니다.

Gartner에서는 PEC(Privacy-Enhancing Computation)을 2021년도에도 중요한 관점으로 생각해 당시에도 "People centricity" 범주로 분류했고, 현재는 “Engineering Trust”의 범주로 분류되어 있습니다.

연구 및 컨설팅 회사에서는 PEC는 데이터 보호를 위한 3가지 중 첫 번째로 중요한 데이터를 처리∙분석할 수 있는 신뢰할 수 있는 환경, 두 번째는 분산화 된 방식으로 처리 데이터와 알고리즘을 처리 및 분석하는 기술, 마지막은 분석 또는 처리 전에 데이터 및 알고리즘을 암호화하는 것으로 구성되어 있다고 말합니다.

첫 번째 환경을 위해서는 신뢰할 수 있는 제 3자와 하드웨어를 신뢰할 수 있는 환경이 있어야 하며 두 번째를 위해선 분산화 된 방식으로 로컬에서 데이터 및 알고리즘을 처리 및 분석하는 기술이 필요하며 마지막을 제공하기 위해서 데이터 분석 이전 단계에서 데이터 및 알고리즘 자체를 암호화하는 기술을 필요합니다.

이제부터 PEC(Privacy-Enhancing Computation)이 왜 필수적이여야 하는지에 대해 이야기하고자 합니다.

앞에서 말한 바와 같이 기업 또는 조직에서 PEC의 기술들의 사용이 요구되고 있습니다.

첫번째로 기업 또는 조직에서 PEC를 구현해야 하는 이유는 소비자의 개인정보에 대한 가능한 위험을 피해야 하기 때문입니다.

사용자들은 자신들의 개인 데이터를 애플리케이션, 웹사이트, 계정 또는 기타 양식에 입력할 때마다 데이터를 비공개로 유지하고, 의도된 목적으로 사용되기를 원합니다.

만약 사용자의 데이터를 보호하기 위한 적절한 테스트 프로세스가 없는 기업이 고객들의 개인정보를 악의적인 공격자에게 노출하게 된다면 해당 기업을 이용하는 고객들은 기업을 자신의 데이터를 노출할 수 있는 위협이 된다 판단하여 해당 기업의 이미지를 자신의 개인적인 데이터를 노출하는 기업으로 판단하여 해당 기업에 대한 신뢰도를 낮게 평가하는 결과를 불러일으킵니다.

이는 기업 또는 조직의 행동에 대한 사람들의 신뢰, 평판, 신뢰 및 신뢰에 영향을 미치므로 기업, 조직은 이러한 사태를 방지하기 위해서 PEC에 대해 기본적으로 구현해야 하는 특징을 지니게 됩니다.

두번째로는 최근 각 정부는 고객 및 소비자의 데이터 보호 및 보호 프로세스를 의무화하고 있습니다.

따라서 각 나라에서 활동중인 기업들은 필수적으로 각 국가에서 지정하는 수준의 PEC를 구현해야 합니다.

다음으로 이러한 특징을 지닌 PEC의 몇가지 주요기술들의 장점, 단점을 이야기하고자 합니다.

일단 데이터 변환단계에서의 주요 기술인 실제 데이터로부터의 엄격한 샘플링, 의미적 접근, 시뮬레이션 시나리오 등 다양한 방법론을 통해 인공적으로 생성된 데이터를 의미하는 기술인 “재현 데이터”는 모델 학습에 필요한 데이터 확보의 어려움을 개선하고 개인 민감정보를 제거함으로 써 데이터 활용성 제고에 기여한다는 장점을 가지지만 특징 데이터 편향을 추가할 위험, 자연적 이상치 누락, 기존 데이터 이외의 새로운 정보를 미제공하는 단점을 제공합니다.

또 기존 암호화 방식과 달리 암호화 상태에서 데이터를 결합하고 연산-분석 등이 가능한 차세대 수학 기법으로, 다양한 계산이 가능하고 양자 내성 암호 안정성을 확보한 ”동형 암호”의 경우 데이터를 명확하게 보지 않고도 정확한 분석이 가능하다는 점, 공동 작업 시 데이터를 노출하지 않고 분석이 가능하며, 사용중인 데이터 보호 요구에 대한 수요를 충족할 수 있는 장점을 가지나 다양한 형태의 동형암호 기술로 복잡성, 느린 처리속도, 전문 인력이 필요하다는 단점을 지닙니다.

마지막으로 데이터셋의 개인 정보에 대한 특정 정보를 유보 및 왜곡하여 제공하는 시스템으로 노이즈 및 매개변수를 추가하는 정확한 수식을 사용하는 “차분 프라이버시”은 개인데이터의 프라이버시를 보장하면서 집계 데이터의 의미 있는 분석을 지원하는 장점과 원본 데이터를 변경하지 않고 데이터를 즉시 변환할 수 있는 장점을 가지지만 복잡한 분석의 어려움과 분석 결과 값의 정확도 손실, 숙련된 전문가가 필요하다는 단점이 있습니다.

그리고 소프트웨어 단계에서의 주요 기술인 기본 정보를 전송하거나 공유할 필요없이 정보가 정확함을 증명할 수 있도록 하는 프라이버시 보호 메시지 프로토콜인 “영지식증명”은 민감데이터를 전송할 필요없이 정보의 유효성을 검증할 수 있다는 장점을 지니나 다양한 작업 그룹 간의 공통 프레임워크를 통한 솔루션 실용화가 필요하다는 단점을 지닙니다.

추가적으로 데이터 및 암호화키를 보호된 상태로 유지하면서, 개인-조직-단말에서 데이터 작업이 가능하도록 하는 분산컴퓨팅 및 암호화 방법을 이야기하는 “다자간 컴퓨팅 알고리즘”은 데이터 기밀을 유지하면서 다자간 통찰력을 공유할 수 있고 사용중인 데이터 보호가 가능해지면서 기계학습 훈련에 실시간으로 사용 가능하며 동형암호보다 빠르다는 장점이 있지만 알고리즘이 대기시간에 민감하여 성능저하 문제가 발생할 수 있다는 단점을 지닙니다.

추가적으로 데이터 샘플을 교환하지 않고 로컬 노드에 있는 다중 데이터 셋으로 기계학습 알고리즘을 훈련할 수 있게 하는 “연합 학습”은 분산 환경 기계학습 알고리즘 훈련이 가능함으로써 데이터 전송 병목현상을 해결하는 장점과 데이터를 로컬에 유지하면서 중앙수준에서 로컬 모델 개선 사항을 공유함으로써 협업 ML을 할 수 있다는 장점을 지니지만 시스템 및 데이터 이질성을 해결하기 위한 표준성이 필요하고 DataOps, ModelOps 배포 및 지속적인 추척/재훈련 전반에 걸쳐 통합하는 인프라 스택이 필요하다는 단점을 지녔습니다.

마지막으로 하드웨어 단계에서의 주요 기술인 하드웨어 기반 TEE를 실행하기 위한 보안 메커니즘인 “기밀 컴퓨팅”은 칩 수준 TEE와 기존 키 관리 및 암호화 프로토콜과 결합하여 데이터 및 IP공유없이 협력이 가능하다는 장점이 있지만 기술적 복잡성과 훈련된 직원이 필요하다는 단점을 지닙니다.

이렇게 PEC의 주요 기술들에 대해 장점과 단점을 보면서 생각한 활용할 방안들은 블록체인이나 개인정보를 요청하는 서비스 상에서 신원을 증명할 때 “영지식증명”을 통한 사용자가 자신의 개인정보를 선택적으로 제공해 자신을 증명하는 서비스나 다양한 인공지능 시스템에서 “연합 학습”을 통한 협업 ML을 통해 개선하거나 외부에 제공하는 통계시스템에 “차분 프라이버시”를 적용해 통계 데이터셋에서 Privacy에 연관된 데이터를 삭제하는 방법으로 사용될 수 있을 듯 합니다.

이러한 활용 방안으로 봤을 때는 PEC기술이 매우 좋을 듯하지만 현실에는 다양한 고려사항이 존재한다는 것을 알 수 있습니다.

첫번째로는 활용방안에서 이야기했던 “영지식증명”에서의 자신임을 증명하기 위해 많은 수식을 계산해야 하는 컴퓨팅 파워 작업을 고려해야함 생각해 볼 수 있었습니다.  
제한적인 데이터를 통해 사용자를 증명하기 위해 많은 수식을 계산하게 되면서 필수적으로제공하는 서비스의 성능이 떨어질 수 있다는 고려사항을 통해 보안과 성능의 관계에 대해 다시한번 더 생각해보는 시간을 가지게 되었습니다.

두번째로는 앞에서 이야기했던 PEC와 더불어 다양한 PEC기술이 과연 현재 서비스 제공 수준에서 동작이 가능한지 고려해 봐야함을 생각해 볼 수 있었습니다.  
만약 현재의 서비스가 영지식증명을 제공할 성능 부족, DataOps, ModelOps 등 CI/CD기능을 가지지 않은 서비스가 갑자기 “영지식증명”, “연합 학습”을 제공하고자 했을 경우 제공하는 서비스에 문제가 생길 수 있다는 점에서 PEC의 기술들과 현재 서비스의 기술들을 분석해 필요한 기능만을 삽입하는 것을 고려해봐야 함을 알 수 있습니다.

마지막으로 이렇게 다양한 PEC 기술을 통해 보호했다 하더라도 해당 방법을 우회하거나 다른 방법으로 사용자의 프라이버시가 공격당할 수 있으므로 공격을 당했을 시의 처리 프로세스나 대응책을 미리 구상해서 가지고 있어야함을 고려해 봐야함을 알 수 있습니다.

이렇게 Gartner가 발표했던 최근의 기술들에 대해 선정이유와 배경 등 다양한 정보들을 차근차근 읽어보며 공부하면서 느낀 점은 현재 제가 스스로 미래에는 어떠한 기술과 보안요소들이 자리잡을 것이라 생각했던 부분에 대해 유사한점과 멀었던 기술들의 목록을 비교해보며 왜 이러한 현상이 발생했는지에 대해 고민해볼 수 있었던 시간이었던 것 같습니다.

또한 미래에는 개인정보나 인공지능, 클라우드, 양자컴퓨팅 등 다양한 분야가 생성하기 전에 보안이 나오는 것이 아니라 발전하고 성숙도가 올라감에 따라 해당 분야에 보안과 연결되게 된 것이라는 점도 역시 추가적 배울 수 있었습니다.

지금까지의 Gartner가 제공하였던 자료와 기술의 발전을 바탕으로 Gartner사에서 어떠한 관점으로 미래를 바라보는지에 대해 생각해보고 고민해봤던 좋은 경험이라 생각합니다.