(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0072770
(43) 공개일자 2019년06월26일

(51) 국제특허분류(Int. Cl.)
H04L 29/06 (2006.01) *H04L 9/08* (2006.01)
(52) CPC특허분류
H04L 63/0428 (2013.01)
H04L 43/16 (2013.01)
(21) 출원번호 10-2017-0173815
(22) 출원일자 2017년12월18일
심사청구일자 2017년12월18일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
홍충선
경기도 용인시 수지구 상현로 30-10 상현마을 성원상떼빌 233-101 (상현동, 상현마을성원상떼빌아파트)
김영기
경기도 수원시 영통구 청명로59번길 45-14, 203호
(74) 대리인
김등용, 김홍석

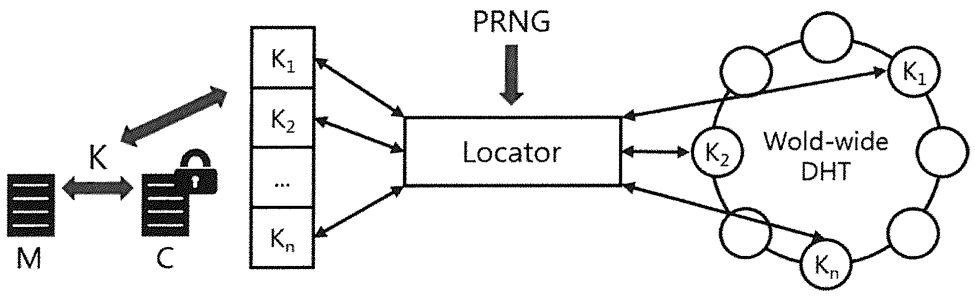
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **강화 학습 기반 암호화 및 복호화 수행 방법 및 이를 수행하는 클라이언트, 서버 시스템**

(57) 요 약

본 발명에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 클라이언트 및 서버 시스템은, 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하는 키 관리 모듈; 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하는 보안 공유(secret sharing) 모듈; 및 상기 임계값을 대한 예측을 수행하는 임계값 예측 모듈을 포함하고, 프라이버시 보호를 위한 Self-Destructing 환경에서 사용자의 요구사항을 만족하는 데이터의 가용성 및 보안성을 향상시킬 수 있다.

대 표 도 - 도1



(52) CPC특허분류
H04L 9/085 (2013.01)
H04L 9/0861 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 2015-0-00274
부처명 미래창조과학부
연구관리전문기관 정보통신기술진흥센터 (IITP)
연구사업명 정보 통신 방송 연구개발사업
연구과제명 ICBMS 플랫폼 간 정보 모델 연동 및 서비스 매쉬업을 위한 스마트 중재 기술 개발
기 여 율 1/1
주관기관 경희대학교 산학협력단
연구기간 2017.03.01 ~ 2018.02.28

명 세 서

청구범위

청구항 1

강화 학습 기반 암호화 및 복호화를 수행하는 클라이언트 및 서버 시스템에 있어서,

데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하는 키 관리 모듈;

상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하는 보안 공유(secret sharing) 모듈; 및

상기 임계값을 대한 예측을 수행하는 임계값 예측 모듈을 포함하는, 클라이언트 및 서버 시스템.

청구항 2

제1 항에 있어서,

사용자로부터 입력 값을 전달받는 UI 모듈; 및

상기 데이터에 대한 암호화 및 복호화를 수행하는 암호화 모듈을 더 포함하는, 클라이언트 및 서버 시스템.

청구항 3

제1 항에 있어서,

상기 임계값 예측 모듈은,

상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 4

제3 항에 있어서,

상기 임계값 예측 모듈은,

상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고,

상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 5

제3 항에 있어서,

상기 임계값 예측 모듈은,

초기 상태 및 액션을 행렬 형태로 구조화하고,

현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하고,

상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 6

제5 항에 있어서,

상기 임계값 예측 모듈은,

상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고,

상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고,

에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고,

상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 7

제6 항에 있어서,

상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는,

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$$ 에 의해 결정되고,

상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복되는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 8

강화 학습 기반 암호화 및 복호화를 수행하는 암호화 및 복호화 방법에 있어서,

데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하는 키 관리 과정;

상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하는 보안 공유(secret sharing) 과정; 및

상기 임계값을 대한 예측을 수행하는 임계값 예측 과정을 포함하는, 암호화 및 복호화 방법.

청구항 9

제8 항에 있어서,

상기 키 관리 과정 이전에,

사용자로부터 입력 값을 전달받는 사용자 인터랙션(UI) 과정; 및

상기 데이터에 대한 암호화 및 복호화를 수행하는 암호화 과정을 더 포함하는, 클라이언트 및 서버 시스템.

청구항 10

제8 항에 있어서,

상기 임계값 예측 과정에서,

상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 11

제10 항에 있어서,

상기 임계값 예측 과정에서,

상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고,

상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 12

제10 항에 있어서,

상기 임계값 예측 과정은,

초기 상태 및 액션을 행렬 형태로 구조화하는 초기 상태 및 액션 행렬 구성 과정;

현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하는 행렬 갱신 과정; 및

상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 임계값 예측 과정을 포함하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 13

제12 항에 있어서,

상기 초기 상태 및 액션 행렬 구성 과정에서, 상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고,

상기 행렬 갱신 과정에서, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고, 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 14

제13 항에 있어서,

상기 행렬 갱신 과정에서, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는,

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$$ 에 의해 결정되고,

상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고,

상기 임계값 예측 과정에서, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신 과정이 반복되는 것을 특징으로 하는, 클라이언트 및 서버 시스템.

청구항 15

강화 학습 기반 암호화 및 복호화를 수행하는 서버에 있어서,

송신 클라이언트 단말로부터 입력 값을 전달받는 사용자 인터랙션(UI) 모듈;

데이터에 대한 암호화 및 복호화를 수행하는 암호화 모듈;

상기 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하고, 상기 임계값을 대한 예측을 수행하는 보안 모듈; 및

분산 해시 테이블(DHT: Distributed Hash Table)에 기반하여, 적어도 하나의 수신 클라이언트 단말과 상기 데이터를 공유하도록 하는 DHT 네트워크 모듈을 포함하는, 서버.

청구항 16

제15항에 있어서,

상기 보안 모듈은,

상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 하는, 서버.

청구항 17

제16항에 있어서,

상기 보안 모듈은,

상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고,

상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 하는, 서버.

청구항 18

제16 항에 있어서,

상기 보안 모듈은,

초기 상태 및 액션을 행렬 형태로 구조화하고,

현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하고,

상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 것을 특징으로 하는, 서버.

청구항 19

제18 항에 있어서,

상기 보안 모듈은,

상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고,

상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고,

에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고,

상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 하는, 서버.

청구항 20

제19 항에 있어서,

상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는,

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$$ 에 의해 결정되고,

상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복되는 것을 특징으로 하는, 서버.

발명의 설명

기 술 분 야

[0001] 본 발명은 암호화 및 복호화 수행 방법에 관한 것이다. 보다 상세하게는, 프라이버시 보호를 위한 강화 학습 기반 암호화 및 복호화 수행 방법 및 이를 수행하는 클라이언트, 서버 시스템에 관한 것이다.

배 경 기 술

[0002] Self-Destructing 기법은 클라우드 컴퓨팅 환경에서 스토리지에 저장되어있는 사용자의 개인정보를 보호하기 위해 2009년 R. Geambasu, T. Kohno, Amit A. Levy, Henry M. Levy에 의해 제안된 시스템으로, 데이터가 사용자로부터 입력받은 특정 시간 이후에 스스로 소멸되도록 한다.

[0003] 이러한 시스템은 파일, 비공개 블로그 게시물, 문서, 이메일, 메시지 등 현대 정보화 사회에 광범위하게 적용될 수 있다. 모든 디지털 컨텐츠의 프라이버시는 데이터를 삭제함으로써 보장할 수 있지만, 제안된 기법은 사용자의 개입이나 별도의 하드웨어가 필요하지 않다는 이점을 갖는다.

[0004] 하지만, 이러한 Self-Destructing 기법을 이용하여 클라우드 컴퓨팅 환경에서 스토리지에 저장되어있는 사용자

의 개인정보를 보호함에 있어, 사용자가 요구하는 시간 이전에 데이터가 삭제될 수 있는 문제가 있다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 대한민국 공개특허 제10-2014-0109337호(2014.09.15) "기본 가치 신호를 이용한 강화학습 방법 및 그 장치"

발명의 내용

해결하려는 과제

[0006] 본 발명의 목적은 프라이버시 보호를 위한 Self-Destructing 기법에 기반한 암호화 및 복호화 방법을 제공하는 데 있다.

[0007] 본 발명의 다른 목적은 암호화 및 복호화에 사용되는 키를 나눌 때 강화학습을 적용하여 데이터의 가용성 및 보안성을 고려한 임계값을 설정하는 데 있다.

과제의 해결 수단

[0008] 전술된 문제점을 해결하기 위한 본 발명에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 클라이언트 및 서버 시스템은, 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하는 키 관리 모듈; 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하는 보안 공유(secret sharing) 모듈; 및 상기 임계값을 대한 예측을 수행하는 임계값 예측 모듈을 포함하고, 프라이버시 보호를 위한 Self-Destructing 환경에서 사용자의 요구사항을 만족하는 데이터의 가용성 및 보안성을 향상시킬 수 있다.

[0009] 일 실시 예에서, 사용자로부터 입력 값을 전달받는 UI 모듈; 및 상기 데이터에 대한 암호화 및 복호화를 수행하는 암호화 모듈을 더 포함할 수 있다.

[0010] 일 실시 예에서, 상기 임계값 예측 모듈은, 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0011] 일 실시 예에서, 상기 임계값 예측 모듈은, 상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0012] 일 실시 예에서, 상기 임계값 예측 모듈은, 초기 상태 및 액션을 행렬 형태로 구조화하고, 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하고, 상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0013] 일 실시 예에서, 상기 임계값 예측 모듈은, 상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고, 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 할 수 있다.

[0014] 일 실시 예에서, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는, $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$ 에 의해 결정될 수 있다. 이때, 상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복되는 것을 특징으로 할 수 있다.

[0015] 본 발명의 다른 측면에 따른, 강화 학습 기반 암호화 및 복호화를 수행하는 암호화 및 복호화 방법은, 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하는 키 관리 과정; 상기 데이터에 대한 유지 시

간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하는 보안 공유(secret sharing) 과정; 및 기 임계값을 대한 예측을 수행하는 임계값 예측 과정을 포함한다.

[0016] 일 실시 예에서, 상기 키 관리 과정 이전에, 사용자로부터 입력 값을 전달받는 사용자 인터랙션(UI) 과정; 및 상기 데이터에 대한 암호화 및 복호화를 수행하는 암호화 과정을 더 포함할 수 있다.

[0017] 일 실시 예에서, 상기 임계값 예측 과정에서, 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0018] 일 실시 예에서, 상기 임계값 예측 과정에서, 상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0019] 일 실시 예에서, 상기 임계값 예측 과정은, 초기 상태 및 액션을 행렬 형태로 구조화하는 초기 상태 및 액션 행렬 구성 과정; 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하는 행렬 갱신 과정; 및 상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 임계값 예측 과정을 포함할 수 있다.

[0020] 일 실시 예에서, 상기 초기 상태 및 액션 행렬 구성 과정에서, 상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고, 상기 행렬 갱신 과정에서, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고, 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 할 수 있다.

[0021] 일 실시 예에서, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는, $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$ 에 의해 결정될 수 있다. 이때, 상기 행렬 갱신 과정에서, 상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 임계값 예측 과정에서, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신 과정이 반복될 수 있다.

[0022] 본 발명의 또 다른 측면에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 서버는, 송신 클라이언트 단말로부터 입력 값을 전달받는 사용자 인터랙션(UI) 모듈; 데이터에 대한 암호화 및 복호화를 수행하는 암호화 모듈; 상기 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하고, 상기 임계값을 대한 예측을 수행하는 보안 모듈; 및 분산 해시 테이블(DHT: Distributed Hash Table)에 기반하여, 적어도 하나의 수신 클라이언트 단말과 상기 데이터를 공유하도록 하는 DHT 네트워크 모듈을 포함한다.

[0023] 일 실시 예에서, 상기 보안 모듈은, 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0024] 일 실시 예에서, 상기 보안 모듈은, 상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0025] 일 실시 예에서, 상기 보안 모듈은, 초기 상태 및 액션을 행렬 형태로 구조화하고, 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하고, 상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하는 것을 특징으로 할 수 있다.

[0026] 일 실시 예에서, 상기 보안 모듈은, 상기 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고, 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택하는 것을 특징으로 할 수 있다.

[0027] 일 실시 예에서, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션

(a')는, $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma Q(s', a') - Q(s, a)]$ 에 의해 결정될 수 있다. 이때, 상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복되는 것을 특징으로 할 수 있다.

발명의 효과

[0028] 본 발명에 따른 강화 학습 기반 암호화 및 복호화 방법은, 프라이버시 보호를 위한 Self-Destructing 환경에서 사용자의 요구사항을 만족하는 데이터의 가용성 및 보안성을 향상시킬 수 있다는 장점이 있다.

[0029] 또한, 본 발명에 따른 강화 학습 기반 암호화 및 복호화 방법은, 사용자의 요구사항과 함께 어플리케이션에 따라 차별적으로 데이터의 가용성 및 보안성을 향상시킬 수 있다는 장점이 있다.

도면의 간단한 설명

[0030] 도 1은 본 발명에 따른 Self-Destructing 기법의 전체적인 시스템 구조를 나타낸다.

도 2는 본 발명에 따른 강화 학습과 관련하여, 상태, 액션 및 보상을 수행하는 개념도를 나타낸다.

도 3은 본 발명에 따른 강화 학습 기반 암호화 및 복호화 방법에서, N개의 같은 키 조각과 이를 복호화 하기 위해 필요한 임계값이 다른 두 그래프를 나타낸다.

도 4는 본 발명에 따른 데이터 암호화 및 복호화를 수행하는 시스템 구조를 나타낸다.

도 5는 본 발명에 따른 강화 학습이 적용된 임계값 예측 과정의 구체적인 알고리즘을 나타낸다.

도 6은 본 발명에 따른 암호화 및 복호화 방법의 흐름도를 나타낸다.

도 7은 본 발명에 따른 임계값 예측 과정의 상세한 흐름도를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0031] 상술한 본 발명의 특징 및 효과는 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명은 다양한 변경을 가할 수 있고 여러 가지 형태를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 본문에 상세하게 설명하고자 한다. 그러나 이는 본 발명을 특정한 개시형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 명세서에서 사용한 용어는 단지 특정한 실시 예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다.

[0032] 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용한다.

[0033] 제1, 제2등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0034] 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는" 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0035] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미가 있다.

[0036] 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않아야 한다.

[0037] 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈", "블록" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다.

[0038] 이하, 본 발명의 바람직한 실시 예를 첨부한 도면을 참조하여 당해 분야에 통상의 지식을 가진 자가 용이하게

실시할 수 있도록 설명한다. 하기에서 본 발명의 실시 예를 설명함에 있어, 관련된 공지의 기능 또는 공지의 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다.

[0039] 이하, 본 발명에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 방법 및 이를 수행하는 클라이언트 및/또는 서버 시스템에 대해 살펴보기로 한다.

[0040] 이와 관련하여, 도 1은 본 발명에 따른 Self-Destructing 기법의 전체적인 시스템 구조를 나타낸다. 한편, 본 발명에 따른 Self-Destructing 기법의 핵심 기술은 분산된 P2P 인프라, 특히 분산 해시 테이블(DHT) 네트워크를 활용하는데 있다. 또한, 분산 해시 테이블 네트워크는 P2P를 구성하는 각 노드들이 Index-Value 쌍을 갖도록 설계되어있다. 한편, 데이터 소멸을 목적으로 하는 Self-Destructing 기법에서 분산 해시 테이블 네트워크를 사용하는 이유는 다음의 세 가지 고유한 특성에 있다.

[0041] - 네트워크를 구성하는 많은 수의 노드들이 다양한 지역에 분산되어 공격자로부터 안전하다.

[0042] - 분산 저장 장치를 위해 설계되어있어 원하는 시간 간격 동안 사용자가 데이터를 사용할 수 있도록 보장한다.

[0043] - 네트워크를 구성하는 노드들이 지속적으로 추가/삭제되며 그 과정에서 데이터가 자동적으로 소멸된다.

[0044] 위와 같은 특성을 지닌 분산 해시 테이블 네트워크를 활용하여 특정 시간 이후에 데이터가 스스로 소멸되도록 함으로써 사용자의 민감 정보가 유출되지 않도록 한다.

[0045] 한편, 도 1을 참조하면, 스템은 데이터를 암호화하기 위해 임의의 암호화 키를 선택하여 암호문을 얻는다. 그리고 Threshold Secret Sharing을 사용하여 암호화 키를 여러 개의 조각으로 나눈다. 이 때, 암호화 키를 얻기 위해 필요한 키 조각의 개수를 임계값으로 정의하며 이 값은 사용자나 애플리케이션에 의해 설정되도록 한다.

[0046] 키 조각의 개수와 임계값이 결정되면 난수기반의 또 다른 임의의 키를 생성한다. 이 임의의 키는 분산 해시 테이블 네트워크에 배포할 키 조각들의 위치를 지정하는데 사용된다. 만약 분산 해시 테이블 네트워크가 여러 종류의 시간제한을 지원한다면 그에 대한 정보를 각 키 조각들에 포함한다. 배포된 키 조각들의 수가 임계값보다 적어질 경우 해당 데이터는 어떤 경우에도 복원할 수 없다.

[0047] 스토리지에 최종적으로 저장되는 데이터 객체는 (임의의 키, 암호문, 전체 키 조각의 개수, 임계값)으로 구성되며, 이때 임의의 키는 분산 해시 테이블 네트워크에 배포할 키 조각들의 위치를 지정하는데 사용된 키를 의미한다. 시스템을 활용하여 생성된 데이터 객체는 이메일 서버를 통해 전송되거나 기존 파일 시스템에 저장될 수 있다. 데이터 객체의 복호화 과정은 암호화 과정의 역순으로 이루어지며 사용자가 설정한 시간이 초과되지 않은 경우에만 가능하다. 먼저, 데이터 객체로부터 임의의 키를 추출하고 키 조각들의 위치를 알아낸다. 확인된 위치를 바탕으로 분산 해시 테이블 네트워크로부터 키 조각들을 얻어낸다. 그리고 얻어낸 키 조각들로 암호화를 수행하는데 사용된 키를 얻어 암호문을 복호화하고 데이터를 얻는다. 위와 같은 과정에서 임계값 보안 공유(Threshold Secret Sharing)에 필요한 임계값은 사용자의 요구사항을 만족하기 위해 중요한 요소라고 할 수 있다.

[0048] 다음으로, 본 발명에 따른 강화 학습(reinforced learning)과 관련하여, 상태, 액션 및 보상을 수행하는 원리에 대해 살펴보기로 한다. 이와 관련하여, 도 2는 본 발명에 따른 강화 학습과 관련하여, 상태, 액션 및 보상을 수행하는 개념도를 나타낸다. 한편, 강화학습은 기계학습의 한 영역으로, 특정 환경에서 정의된 에이전트(10)가 현재의 상태를 인식하여 선택 가능한 행동들 중에서 보상을 최대화하는 액션(action)을 선택하여 문제를 해결하는 방법이다. 이때, 에이전트(10)가 얻게 되는 보상(reward)은 양수와 음수 둘 다 가능하며, 알고리즘을 바탕으로 에이전트(10)가 앞으로 얻게 될 누적 보상을 최대화하는 일련의 행동을 정의하는 방법이다. 도 2에 도시된 바와 같이, 에이전트(10)는 환경(environment)(20)의 제약 조건하에서, 가능한 상태(state)들 중 특정 상태를 선택하고, 이에 기반하여 액션(action)을 수행한다. 이러한 선택된 상태와 수행된 액션을 반복함에 따라 보상(reward)이 변하게 되고, 특정 조건을 만족할때까지 에이전트(10)는 일련의 과정을 반복한다.

[0049] 한편, 본 발명에서 수행되는 강화 학습 기반 암호화 및 복호화 방법은 이러한 상태, 및 액션의 선택 및 수행이 반복되어 이에 따라 보상이 산출되므로, SARSA(State-Action-Reward-State-Action)라고 지칭될 수 있다. 즉, 본 발명에서는 강화학습의 여러 종류 중 SARSA(State-Action-Reward-State-Action)를 이용하여 문제를 해결하는 방법을 제안한다.

[0050] 한편, 전술한 바와 같이, 호화 키를 얻기 위해서 필요한 최소한의 키 조각의 개수를 임계값으로 정의하며 이는 데이터의 가용성 및 보안성과 관련되어 있다. 이와 관련하여, 도 3은 본 발명에 따른 강화 학습 기반 암호화 및

복호화 방법에서, N개의 같은 키 조각과 이를 복호화 하기 위해 필요한 임계값이 다른 두 그래프를 나타낸다. 임계값의 비율이 95%인 경우에는 데이터가 유지되어야 하는 시간을 충족하지 못했으며, 45%인 경우에는 일정 기간이 지난 뒤에도 키의 조각이 완전히 사라지지 않았으므로 보안상의 문제를 야기할 수 있다. 이러한 문제는 노드들이 지속적으로 추가/삭제되는 분산 해시 테이블 네트워크의 특수성 때문이다.

[0051] 한편, 도 3에서 데이터의 가용성 및 보안성을 모두 고려하기 위해서는 이상적인 그래프와 가장 유사한 결과를 갖는 임계값을 찾아야 한다. 따라서 본 발명에서는 이를 해결하기 위해 그래프의 유사도와 임계값을 바탕으로 강화학습을 적용하여 최적의 임계값을 찾는 방법을 제안한다.

[0052] 한편, 도 1 내지 도 3에서의 개념에 따라 데이터의 가용성 및 보안성을 고려한 강화 학습 기반 암호화 및 복호화를 수행하기 위한 시스템 구조는 도 4와 같다. 즉, 도 4는 본 발명에 따른 데이터 암호화 및 복호화를 수행하는 시스템 구조를 나타낸다. 즉, 본 발명에서는 프라이버시 보호를 위한 Self-Destructing 환경에서 강화학습의 한 종류인 SARSA를 적용하여 데이터의 가용성 및 보안성을 고려한 임계값을 결정하는 기술을 제안한다.

[0053] 도 4를 참조하면, 강화 학습 기반 암호화 및 복호화를 수행하는 클라이언트 및 서버 시스템(1000)은 클라이언트(100), 암호화/복호화 부(200), 및 DHT 네트워크상의 복수의 클라이언트들(300)을 포함한다. 이와 관련하여, 강화 학습 기반 암호화 및 복호화를 수행하는 암호화/복호화 부(200)는 서버에 해당할 수 있지만, 이에 한정되는 것은 아니고, 응용(application)에 따라 다양하게 변형되어 사용될 수 있다. 예를 들어, 사용자 단말에 해당하는 클라이언트(100) 내에 강화 학습 기반 암호화 및 복호화를 수행하는 일부 구성이 포함되고, 나머지 구성은 서버에 의해 이루어질 수 있다. 대안적으로, 모든 암호화 및 복호화가 클라이언트(100) 및 복수의 클라이언트들(300) 중 해당 데이터(컨텐츠)를 수신하려는 수신 클라이언트에 의해 이루어질 수 있다.

[0054] 한편, 암호화/복호화 부(200)는 사용자 인터랙션(UI: User Interaction) 모듈(210), 암호화 모듈(Cryptography Module, 220), 보안(Secret) 모듈(250) 및 DHT 네트워크 모듈(DHT Network Module, 260)을 포함할 수 있다. 한편, 보안 모듈(250)은 키 관리 모듈(Key Management Module, 251), 보안 공유 모듈(Secret Sharing Module, 252), 및 임계값 예측 모듈(Threshold Estimation Module, 253)을 포함하도록 구성된다.

[0055] UI 모듈(210)은 사용자로부터 입력 값을 전달받도록 구성된다. 암호화 모듈(220)은 데이터에 대한 암호화 및 복호화를 수행하도록 구성된다. 한편, 키 관리 모듈(251)은 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하도록 구성된다. 또한, 보안 공유 모듈(252)은 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하도록 구성된다. 또한, 임계값 예측 모듈(253)은 이러한 임계값을 대한 예측을 수행하도록 구성된다. 한편, DHT 네트워크 모듈(260)은 분산 해시 테이블(DHT: Distributed Hash Table)에 기반하여, 적어도 하나의 수신 클라이언트 단말(200)과 데이터를 공유하도록 한다.

[0056] 전술된 임계값 예측 모듈(253)의 상세 동작에 대해 살펴보면 다음과 같다. 이와 관련하여, 제안하는 강화학습을 적용한 임계값 결정은 임계값 예측 모듈(253)에서 다음과 같은 과정을 통해 수행된다. 즉, 임계값 예측 모듈(253)은 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 임계값을 예측할 수 있다. 이와 관련하여, 표 1은 본 발명에 따른 강화 학습을 위한 매개변수와 이에 따른 상태, 액션 및 보상을 나타낸다.

표 1

[0057]

State	N, T
Action	Select N, T
Reward	Similarity with Ideal Graph

[0058] 표 1에 표현된 바와 같이, 상태(state)는 복수의 서로 다른 상태인 N, T를 포함할 수 있고, 이에 따라 액션(action)은 상태 N, T 중 어느 하나를 선택(select)하는 것에 해당한다. 한편, 이러한 상태와 선택된 액션에 대한 보상(reward)은 도 2에 표시된 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서 이상적(ideal) 형태에 가장 근접하도록 선택될 수 있다.

[0059] 한편, SARSA 알고리즘을 적용한 임계값 예측에 필요한 매개변수는 State, Action, Reward이며, 각각은 전체 키 조각의 개수 및 임계값, 전체 키 조각의 개수 및 임계값을 선택하는 행위, 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 측정된 이상적인 데이터 가용성 그래프와의 유사도로 정의할 수 있다. 학습과정이 시작되면 초기 State 및 Action을 행렬형태로 구조화하고 현재 State 및 Action, 다음 State 및 Action과 Reward를 바탕으로 행렬을 갱신한다. 이러한 과정을 통해 보상을 최대화하는 전체 키 조각의 개수와 임계값을 찾아 궁극적으로 사용자가 Self-Destructing 시스템을 사용하는데 있어 원하는 시간 동안만 데이터를 사용할 수 있도록 하며, 그 이후에는 데이터가 스스로 소멸될 수 있도록 한다.

[0060] 전술된 동작들을 수행하는 임계값 예측 모듈(253)에서의 동작에 대해 설명하면 다음과 같다. 즉, 임계값 예측 모듈(253)은 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 보상이 최대화되도록 상기 임계값을 예측할 수 있다. 또한, 임계값 예측 모듈(253)은 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측할 수 있다.

[0061] 이러한 임계값 예측을 위한 구체적인 알고리즘에 대해 살펴보면 다음과 같다. 이와 관련하여, 도 5는 본 발명에 따른 강화 학습이 적용된 임계값 예측 과정의 구체적인 알고리즘을 나타낸다. 도 5에 표현된 알고리즘과 관련하여, 임계값 예측 모듈(253)은 초기 상태 및 액션을 행렬 형태로 구조화하고, 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신하도록 구성된다. 또한, 임계값 예측 모듈(253)은 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측하도록 구성될 수 있다.

[0062] 구체적으로, 임계값 예측 모듈(253)은 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택할 수 있다. 또한, 임계값 예측 모듈(253)은 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택할 수 있다. 이와 관련하여, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는, $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$ 에 의해 결정될 수 있다. 한편, 상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복될 수 있다.

[0063] 다음으로, 본 발명의 다른 양상에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 서버에 대해 살펴보기로 하자. 이와 관련하여, 전술한 바와 같이, 도 4에서 암호화/복호화부(200)가 본 발명에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 서버에 해당할 수 있다. 서버(200)는 사용자 인터랙션(UI) 모듈(210), 암호화 모듈(220), 보안 모듈(250), 및 DHT 네트워크 모듈(260)을 포함한다.

[0064] 사용자 인터랙션(UI) 모듈(210)은 송신 클라이언트 단말(100)로부터 입력 값을 전달받도록 구성된다. 암호화 모듈(220)은 데이터에 대한 암호화 및 복호화를 수행하도록 구성된다. 보안 모듈(250)은 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리하고, 상기 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행하고, 상기 임계값을 대한 예측을 수행하도록 구성된다. 또한, DHT 네트워크 모듈(260)은 분산 해시 테이블(DHT: Distributed Hash Table)에 기반하여, 적어도 하나의 수신 클라이언트 단말(200)과 상기 데이터를 공유하도록 구성된다.

[0065] 이때, 보안 모듈(250)은 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측할 수 있다. 이와 관련하여, 보안 모듈(250)은 상기 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측할 수 있다. 또한, 보안 모듈(250)은 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측할 수 있다.

[0066] 구체적으로, 보안 모듈(250)은 초기 상태 및 액션을 행렬 형태로 구조화하고, 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신할 수 있다. 또한, 보안 모듈(250)은 상기 보상이 최대화된 것으로 판단되면, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측할 수 있다.

[0067] 구체적으로, 보안 모듈(250)은 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성하고, 상기 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택할 수 있다. 또한, 보안 모듈(250)은 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택할 수

있다. 이때, 상기 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는, 전술한 바와 같이 $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$ 에 의해 결정될 수 있다. 이와 관련하여, 상기 제2상태(s') 및 상기 제2액션(a')를 상기 제1상태(s) 및 상기 제1액션(a)으로 대체하고, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신이 반복될 수 있다.

[0068] 한편, 본 발명의 또 다른 양상에 따른 강화 학습 기반 암호화 및 복호화를 수행하는 암호화 및 복호화 방법에 대해 살펴보기로 하자. 이와 관련하여, 도 6은 본 발명에 따른 암호화 및 복호화 방법의 흐름도를 나타낸다. 이때, 각 과정은 순차적으로 수행되는 것으로 도시되었지만, 응용에 따라 다양하게 그 순서가 변경되거나 또는 하나 이상의 과정들이 동시에 병렬적으로 수행될 수 있다. 도 6을 참조하면, 암호화 및 복호화 방법은 사용자 인터랙션(UI) 과정(S100), 암호화 과정(S150), 키 관리 과정(S200), 보안 공유(secret sharing) 과정(S300), 및 임계값 예측 과정(S400)을 포함한다.

[0069] 사용자 인터랙션(UI) 과정(S100)에서, 사용자로부터 입력 값을 전달받는다. 암호화 과정(S150)에서, 데이터에 대한 암호화 및 복호화를 수행한다. 한편, 키 관리 과정(S200)에서, 데이터에 대한 암호화 및 복호화를 수행하는데 필요한 암호화 키를 관리한다. 또한, 보안 공유 과정(S300)에서, 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 임계값에 대한 보안 공유를 수행한다. 또한, 임계값 예측 과정(S400)에서, 상기 임계값을 대한 예측을 수행한다.

[0070] 한편, 본 발명에 따른 강화 학습 기반 암호화 및 복호화 방법 중 임계값 예측 과정(S400)에 대해 상세하게 살펴보면 다음과 같다. 이와 관련하여, 임계값 예측 과정(S400)에서, 상태(state), 액션(action), 및 보상(reward)을 포함하는 매개변수에 의해, 상기 보상이 최대화되도록 상기 임계값을 예측할 수 있다. 구체적으로, 임계값 예측 과정(S400)에서, 암호화 키를 얻기 위해 필요한 전체 키 조각의 개수 및 상기 임계값을 선택하고, 상기 선택된 전체 키 조각의 개수 및 임계값을 바탕으로 상기 보상이 최대화되도록 상기 임계값을 예측한다. 또한, 임계값 예측 과정(S400)에서, 데이터에 대한 유지 시간(lifetime)과 가용성(availability)에 대한 그래프에서, 상기 보상이 최대화되는 것과 연관하여 가장 이상적 형태의 그래프에 가장 유사한 형태가 되도록 상기 임계값을 예측한다.

[0071] 한편, 도 7은 본 발명에 따른 임계값 예측 과정의 상세한 흐름도를 도시한다. 도 7을 참조하면, 임계값 예측 과정은 초기 상태 및 액션 행렬 구성 과정(S410), 행렬 갱신 과정(S420), 보상 판단 과정(S430), 및 임계값 예측 과정(S440)을 포함한다.

[0072] 초기 상태 및 액션 행렬 구성 과정(S410)에서, 초기 상태 및 액션을 행렬 형태로 구조화한다. 또한, 행렬 갱신 과정(S420)에서, 현재 상태 및 액션과 다음 상태 및 액션을 바탕으로 상기 행렬을 갱신한다. 한편, 보상 판단 과정(S430)에서, 가장 최근 갱신된 행렬에 기반하여, 보상이 최대화된 것으로 판단할 수 있는지를 결정한다. 이때, 특정 조건을 만족하여 상기 보상이 최대화된 것으로 판단되면, 임계값 예측 과정(S440)에서, 가장 최근의 상태, 액션 및 보상에 기반하여 상기 임계값을 예측할 수 있다. 반면에, 특정 조건을 만족하지 못하여 상기 보상이 최대화되지 않았다고 판단되면, 행렬 갱신 과정(S420)을 반복할 수 있다.

[0073] 구체적으로, 초기 상태 및 액션 행렬 구성 과정(S410)에서, 초기 상태 및 액션에 따라 행렬 Q(s,a)를 구성한다. 또한, 행렬 갱신 과정(S420)에서, 행렬 Q(s,a)을 이용하여 제1상태(s)로부터 제1액션(a)을 선택하고, 에피소드의 각 스텝에서 상기 제1액션(a)을 선택하고, 제1보상(r)과 제2상태(s')를 관측하고, 상기 제2상태(s')로부터 제2액션(a')을 선택하다. 또한, 행렬 갱신 과정(S420)에서, 제1상태(s) 및 상기 제1액션(a)과 상기 제2상태(s') 및 상기 제2액션(a')는, $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]$ 에 의해 결정될 수 있다. 한편, 임계값 예측 과정(S440)에서, 상기 대체된 제1상태(s)에 대응하는 보상이 특정 조건을 만족할 때까지 상기 행렬 갱신 과정(S420)이 반복될 수 있다.

[0074] 이상에서는 본 발명에 따른 강화 학습 기반 암호화 및 복호화 방법 및 이를 수행하는 클라이언트 및/또는 서버 시스템에 대해 살펴보았다. 이와 관련하여, 클라이언트 및 서버 시스템, 암호화 및 복호화 방법 및 강화 학습 기반 암호화 및 복호화를 수행하는 서버에서 각각 설명한 내용들은 상호 결합되어 이용될 수 있음은 물론이다.

[0075] 한편, 전술된 강화 학습 기반 암호화 및 복호화 방법 및 이를 수행하는 클라이언트 및/또는 서버 시스템은, 클라우드 스토리지에 저장된 사용자의 개인 정보를 보호하기 위한 방안으로 적용할 수 있다. 또한, 파일, SNS 게시물, 문서, 이메일, 메시지 등의 모든 디지털 컨텐츠에 광범위하게 적용할 수 있다. 뿐만 아니라, 다양한 분야

에서 임계값 예측을 위해 강화학습을 적용하는 사례로 활용될 수 있다.

[0076] 한편, 본 발명에서 제안하는 해결방안을 바탕으로 애플리케이션, 클라우드 시스템 등의 서비스를 제공하는 사업에 적용 가능하다.

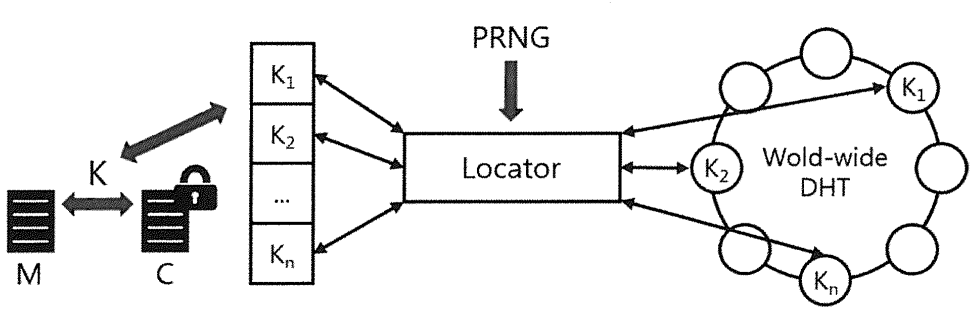
[0077] 본 발명의 적어도 일 실시 예에 따른 강화 학습 기반 암호화 및 복호화 방법은, 프라이버시 보호를 위한 Self-Destructing 환경에서 사용자의 요구사항을 만족하는 데이터의 가용성 및 보안성을 향상시킬 수 있다는 장점이 있다.

[0078] 또한, 본 발명의 적어도 일 실시 예에 따른 강화 학습 기반 암호화 및 복호화 방법은, 사용자의 요구사항과 함께 어플리케이션에 따라 차별적으로 데이터의 가용성 및 보안성을 향상시킬 수 있다는 장점이 있다.

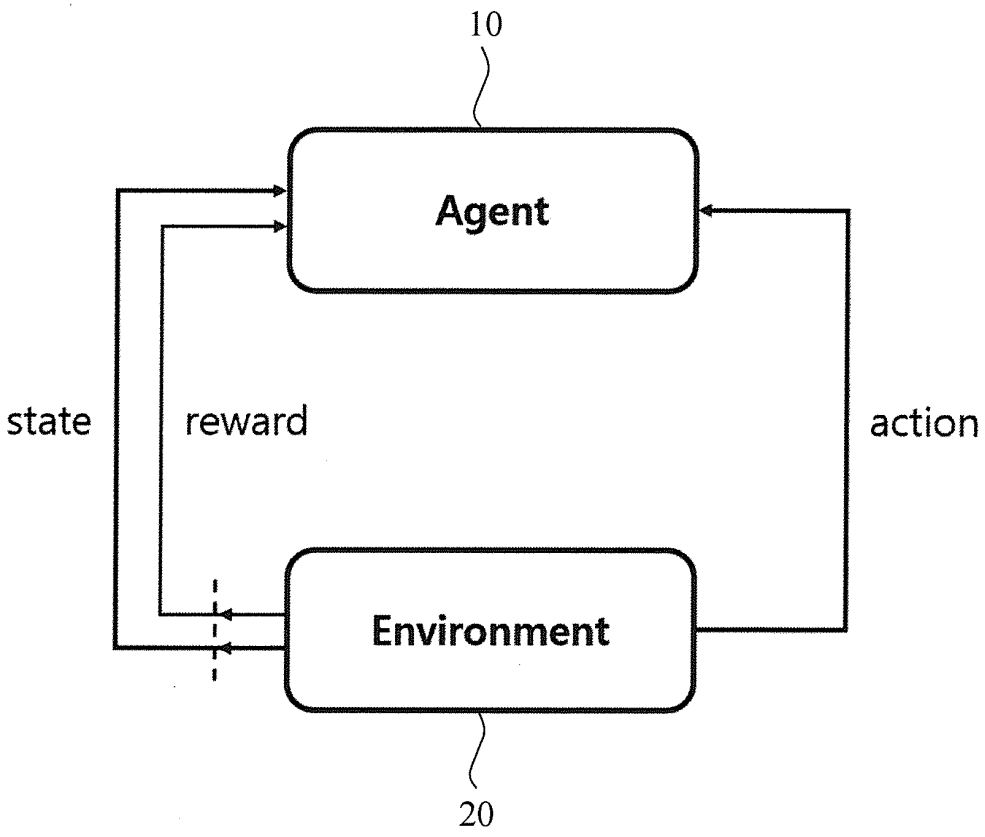
[0079] 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능뿐만 아니라 각각의 구성 요소들은 별도의 소프트웨어 모듈로도 구현될 수 있다. 상기 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다. 적절한 프로그램 언어로 쓰여진 소프트웨어 어플리케이션으로 소프트웨어 코드가 구현될 수 있다. 상기 소프트웨어 코드는 메모리에 저장되고, 제어부(controller) 또는 프로세서(processor)에 의해 실행될 수 있다.

도면

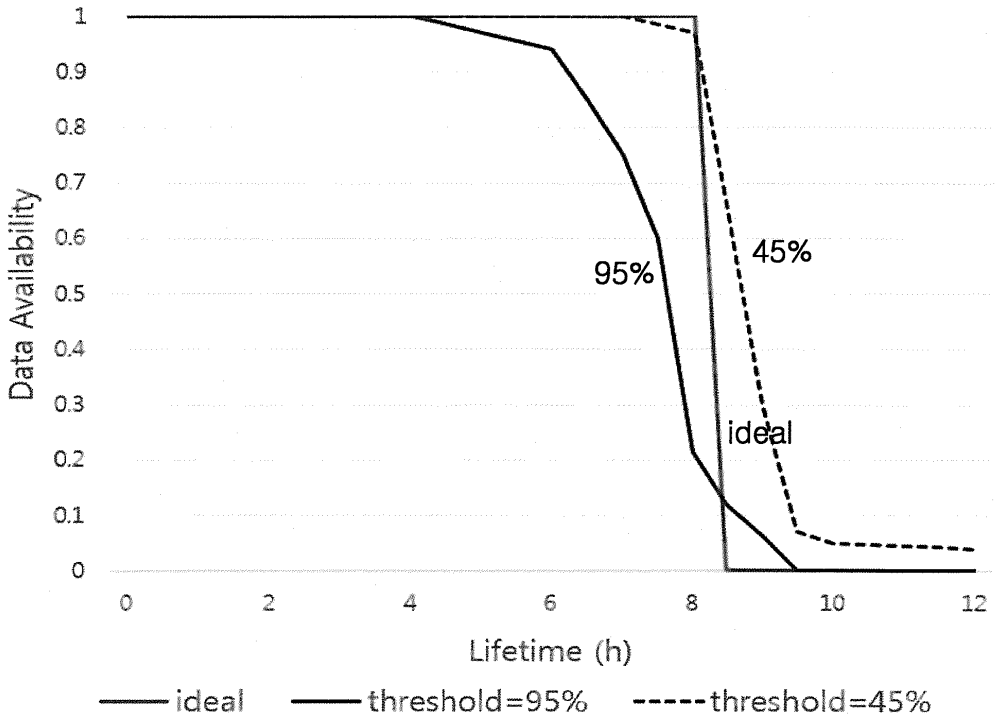
도면1



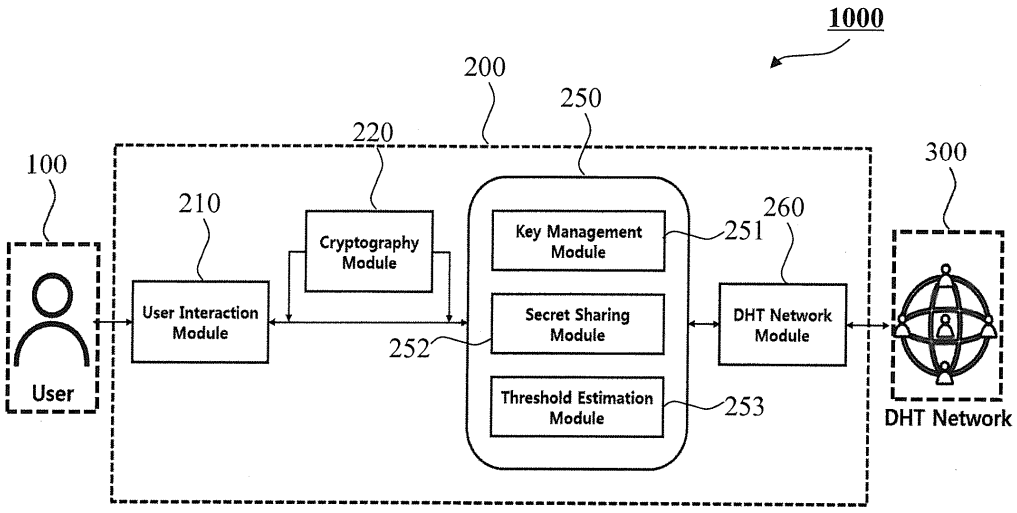
도면2


10
Agent
state
reward
action
Environment
20

도면3


Data Availability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
Lifetime (h)
95%
45%
ideal
ideal
threshold=95%
threshold=45%

도면4

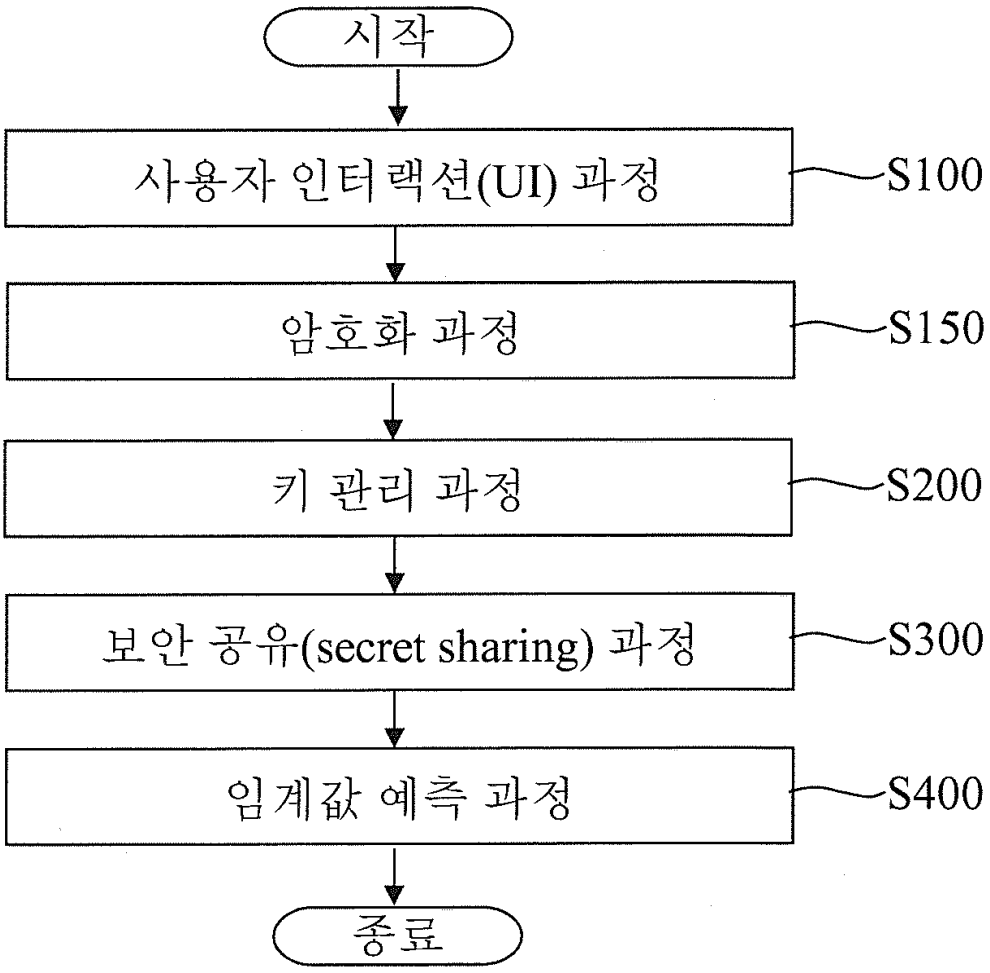


도면5

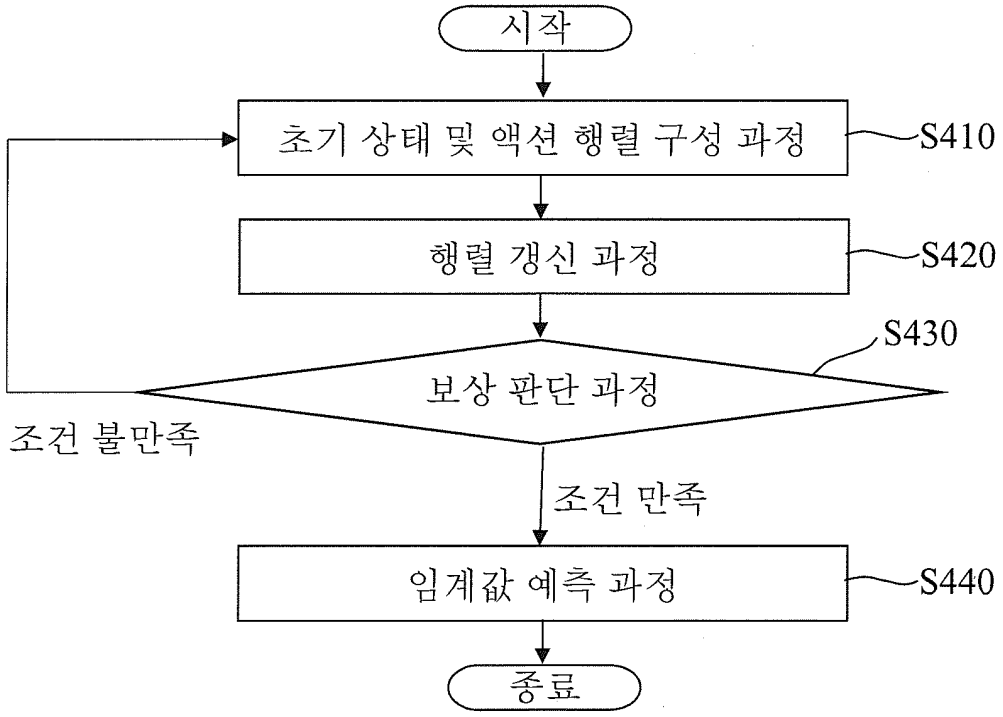
Algorithm : Reinforcement Learning Based Threshold Estimation

1: Initialize $Q(s, a)$
2: **Repeat** (for each episode)
3: Initialize s
4: Choose a from s using Q (ε-greedy)
5: **Repeat** (for each step of episode)
6: Take action a, observe r, s'
7: Choose a' from s' using Q (ε-greedy)
8: $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma Q(s', a') - Q(s, a)]$
9: $s \leftarrow s'; a \leftarrow a'$
10: **Until** s is terminal

도면6


시작
사용자 인터랙션(UI) 과정
S100
암호화 과정
S150
키 관리 과정
S200
보안 공유(secret sharing) 과정
S300
임계값 예측 과정
S400
종료

도면7


시작
초기 상태 및 액션 행렬 구성 과정
S410
행렬 갱신 과정
S420
S430
보상 판단 과정
조건 불만족
조건 만족
임계값 예측 과정
S440
종료