

# キャッシュを包括するウェブセキュリティモデル

島本 隼人<sup>1</sup> 矢内 直人<sup>1</sup> 岡村 真吾<sup>2</sup> 藤原 融<sup>1</sup>

**概要:** ウェブの安全性解析として形式手法を用いた手法が提案されている。しかし、この既存研究のモデルには現実のウェブでは欠かせない概念であるキャッシュが含まれていない。実際に Browser Cache Poisoning 攻撃といったキャッシュを用いた HTTPS に対する攻撃法が存在することから、キャッシュの概念はウェブの安全性解析には不可欠である。本研究では、キャッシュを包括したウェブセキュリティモデルを提案する。提案モデルではキャッシュの概念に加え、その運用を行う HTTP ヘッダや中継者の概念も新たに提案する。

**キーワード:** システム, キャッシュ, ウェブセキュリティモデル, セキュリティモデル, 形式手法

## Toward Formal Verification of Web Security with Cache

HAYATO SHIMAMOTO<sup>1</sup> NAOTO YANAI<sup>1</sup> SHINGO OKAMURA<sup>2</sup> TORU FUJIWARA<sup>1</sup>

**Abstract:** Cache is an important concept in a web system in order to provide them efficiently. Since a browser cache poisoning attack, where an adversary utilizes a web cache, has been proposed recently, analyzing the security under environments with cache is necessary. In this work, we propose a new security model including the cache. We also implement several parts of the model in Alloy.

**Keywords:** system, cache, web security model, security model, formal method

### 1. はじめに

#### 1.1 研究背景

ウェブ上で個人情報を取り扱う機会の増加にともない、ウェブ通信の解析がより重要となっている。ウェブの解析はウェブモデルを用いた形式手法により自動化できることから、これらの研究が近年盛んである。しかし、もしモデルが包括している機能や要件定義が不十分である場合、適切な解析が行えない。そのため、使用するウェブモデルが適切であることが重要となる。

一般に、安全性要件の解析に使われるモデルはセキュリティモデルと呼ばれる。セキュリティモデルでは攻撃者の能力、運用する環境などの条件、許容できる被害レベルなどの安全性が定義されており、攻撃者の能力など各パラメー

タを数値化することで定量的な評価を可能とする。直観的には満たすべきセキュリティモデルを選択し、実装システムがそのモデルに含まれるかどうかを考察することで、実装モデルの安全性が確認できる。また、この解析を形式手法ツールを用い自動化することにより、安全性解析にかかる労力を大幅に削減できる。

Ahkawe ら<sup>?</sup>により提案されたモデルは、Web のセキュリティ機能を導入した初のモデルであり、このモデルに基づいた安全性解析が多くの論文で使用されている。

(引用つける)

しかしながら本研究では Ahkawe らのモデルにはキャッシュの概念が含まれていないことを指摘する。現在、キャッシュはその利便性から広く使われている技術であり、ウェブの安全性を考える上で欠かすことができない。事実、キャッシュを利用した攻撃法も近年の成果において報告されている<sup>?</sup>。以上のことから本研究はキャッシュの動作を包括するウェブセキュリティモデルの提案を目的とする。

<sup>1</sup> 大阪大学  
Osaka University

<sup>2</sup> 奈良工業高等専門学校  
National Institute of Technology, Nara College

## 1.2 貢献

本研究で提案するウェブセキュリティモデルはウェブで使用するキャッシュの動作を包括している。また、ウェブ上でキャッシュを搭載しうる存在としてクライアントとウェブサーバに加えて中継者があり、提案モデルではキャッシュの動作に必要な要素として導入している。さらに、キャッシュを実際に利用する際に必要となる HTTP ヘッダも新たに導入している。例えば、キャッシュの動作の指定に必要な Cache-Control ヘッダや有効期限の計算に必要なレスポンスの生成時刻を記述する Date ヘッダである。

(記述中)

## 1.3 関連研究

Ahkawe らによる既存モデル<sup>?</sup>は本研究で提案するウェブセキュリティモデルのベースとなっている。このモデルは Bau らによる昨今の形式手法を用いた研究の調査論文<sup>?</sup>においても取り上げられており、現在のウェブ環境においてよく使用される HTTP や DNS といった概念を包括し、現在のウェブの環境に最適なものである。しかしこのモデルにはキャッシュが包括されていない。この既存モデルの詳細は??節に記述する。また、キャッシュを用いた攻撃法としてブラウザキャッシュ汚染攻撃<sup>?</sup>が報告されており、提案モデルの能力評価に使用する。この攻撃法の詳細についても??節に記述する。加えて、SSL/TLS の安全性についての研究<sup>?</sup>があり、この研究で想定されている攻撃者の能力を提案モデルが含む脅威モデルとして利用している。

また、形式手法を用いた研究として、耐タンパー性を持つソフトウェアの動作を行うハードウェアの検証<sup>?</sup>があり、形式手法が脆弱性とその対策に有効であることが示されている。同様に、OS カーネルの検証<sup>?</sup>が形式手法を用いて行われており、形式手法を安全性証明として利用できることを示している。さらに、形式手法は AndroidOS の検証<sup>?</sup>にも利用されている。この研究は OS のユーザ権限に着目した安全性検証を目指している。

(記述中)

## 2. Ahkawe らのモデルの問題

### 2.1 モデルの概要

Ahkawe らによる既存モデル<sup>?</sup>に包括されている概念は大きく三つである。一つ目はブラウザである。ブラウザはクライアントとしての基本的な動作と Cookie やスクリプトなどの動作が想定されている。二つ目はウェブサーバである。ウェブサーバでは HTTP レスポンスに対応する動作と DNS に対応する動作の二点に注目している。また、ネットワークの末端に位置し IP アドレスを保有するしていることが想定されている。三つ目はネットワークである。ネットワークには通信プロトコルや、それに必要な要素が包括

される。既存モデルでは HTTP や HTTPS、DNS が包括されており、これらの運用に必要となる HTTP ヘッダや認証局 (CA) など含まれる。

また、既存モデルでは三種類の脅威モデルが想定されている。一つ目は web attacker と呼ばれる攻撃者である。web attacker は最も基本的な攻撃能力を持つ攻撃者であり、以下の能力を有する。

- 任意の HTTP リクエストの送信
- 管理するウェブサーバでの任意のレスポンスの生成
- 所有するドメインの正当なサーバ証明書の生成

二つ目の network attacker は web attacker の能力に加えて暗号化されていない通信の盗聴と改ざん、自己署名サーバ証明書を用いたなりすましの能力を持つ。三つ目の gadget attacker は web attacker の能力に加えてハイパーリンクや画像といった特定の形式のコンテンツを通信内容に挿入できる能力を持つ。

### 2.2 モデルの潜在的脆弱性

Ahkawe らによる既存モデル<sup>?</sup>では、キャッシュの概念が包括されていない。キャッシュはウェブサーバからクライアントへ送信されたレスポンスを保存する領域を持ち、そのレスポンスの再利用によって通信時間を短縮でき通信の効率化をおこなえる。このような利点からキャッシュは現在広く使用されている。

一方で、キャッシュを利用した攻撃法としてブラウザキャッシュ汚染攻撃 (BCP 攻撃)<sup>?</sup>を紹介する。BCP 攻撃は中間者攻撃の一つであり、ウェブサーバから送信されるレスポンスを中間者攻撃によって改ざんし、ブラウザに搭載されているキャッシュに保存させる。これにより、改ざんレスポンスがキャッシュからリロードされるため、ブラウザに悪影響を与えることが可能になる。たとえば、レスポンス内に悪質なサイトのリンクを追加することでブラウザを誘導したり、スクリプトを記述することによりブラウザに任意の動作を強制することができる。この攻撃法の特徴はキャッシュによるレスポンスの再利用機能を利用することにより、一度の中間者攻撃で長期に渡る悪影響が可能になる点である。また、この攻撃法は HTTPS に対する攻撃も可能である。この場合には、攻撃者は中間者攻撃を行う際に自己署名証明書を使用して正当な接続先になりすます必要がある。このなりすましに使用する自己署名証明書の安全性をブラウザでは確認できないため、ブラウザではユーザに対して認証警告が表示する。この認証警告をユーザが無視することが HTTPS に対する BCP 攻撃の条件である。

この BCP 攻撃のようにキャッシュを利用した攻撃法が存在すること、また、加えてキャッシュは一般に使用されているものであることから、ウェブの安全性解析を行う上でキャッシュは必要不可欠な概念である。したがって、

キャッシュを導入したモデルを提案する必要がある。

## 2.3 キャッシュを包括するセキュリティモデルにむけて

キャッシュを導入したモデルの提案に伴って、合わせて必要となる概念がある。まず、中継者と呼ばれる概念である。HTTP において、キャッシュはクライアントとウェブサーバの他に中継者に搭載される。ここで中継者はネットワークの末端のクライアントとウェブサーバの経路上に存在するものであり、具体的にはプロキシやゲートウェイを指す。プロキシは通信経路を最適化することで通信効率を向上させ、また、クライアントの匿名性を保持するために使用される。ゲートウェイは内部のネットワークとウェブ（外部ネットワーク）を接続するサーバであり、ゲートウェイで翻訳を行うことにより異なるプロトコルの通信を可能にする。これらもキャッシュと同じくウェブで広く使用されている要素であるとともに、またキャッシュを使用する概念であることから、キャッシュを導入したモデルを提案する際にはこの中継者の概念も追加する必要がある。モデルにはこれらに加えて、キャッシュおよび中継者の動作に使用されるヘッダが新たに必要となる。レスポンスをキャッシュに保存する際に適応するオプションやキャッシュ内での有効期限はレスポンス内のヘッダで指定する。このヘッダはキャッシュの運用にのみ使用されるヘッダであり、提案モデルには導入しておく必要がある。同様に、中継者の動作に必要なヘッダについても追加する必要がある。

## 3. 提案モデルの設計

本研究で提案するウェブセキュリティモデルの特徴はキャッシュ、ヘッダ、中継者の大きく三つに分けて考えることができる。

### 3.1 キャッシュのモデル定義

提案モデルではキャッシュの基本的な運用に対する動作を包括する。ここで包括する動作は以下の三つに分類できる。

#### 3.1.1 レスポンスの格納と再利用

キャッシュは以下の条件すべてを満たしている場合にレスポンスを保存することができる。

- キャッシュを禁止するオプションが含まれていない
- 有効期限を計算可能、もしくは、キャッシュを許可する記述がある

また、リクエストが要求している URI をもつレスポンスを保持しているキャッシュは、以下のいずれかに該当する場合にそのレスポンスを再利用してリクエストに返送することができる。

- レスポンスが有効期限内
- 有効期限を過ぎているレスポンスの再利用が許可され

ている

- レスポンスが検証済み

#### 3.1.2 キャッシュの検証

キャッシュにおける検証とは、キャッシュ内のレスポンスに含まれているコンテンツがオリジンサーバ内の最新のコンテンツと一致するかを確認する動作である。検証が必要になった場合、キャッシュはオリジンサーバに対して“条件付きリクエスト”を送信する。条件付きリクエストとはヘッダで指定した条件を満たしている場合に完全なレスポンスを返すリクエストである。この条件にはコンテンツの最終更新時間、もしくは、レスポンス内の Etag ヘッダから得られる固有値を使用する。最終更新時刻を用いる場合には If-Modified-Since ヘッダを、固有値を使用する場合には If-None-Match ヘッダを使用した条件付きリクエストを生成する。通信において問題が発生しなかった場合、このリクエストに対するレスポンスの状態コードは 200(OK) もしくは 304(Not Modified) が想定される。200(OK) の場合、キャッシュはキャッシュ内のレスポンスを破棄し、新たに取得したレスポンスを使用する。このとき格納可能条件を満たしていれば、新しいレスポンスをキャッシュしてもよい。304(Not Modified) であった場合、キャッシュ内のレスポンスのヘッダを今取得したレスポンスのヘッダに置き換えて再利用する。このヘッダの置換によりレスポンスの有効期限を延長する。このヘッダの置換時に Waring ヘッダが存在する場合、1xx および 2xx の値は除去する。また、同じ URI に対するレスポンスが同じキャッシュに複数存在する場合、検証結果として送信されるレスポンスに記述されている検証子に該当するレスポンスのヘッダを置換して再利用する。複数のレスポンスに該当する場合には、そのレスポンスのうち最新のレスポンスを採用する。

#### 3.1.3 キャッシュの有効期限

最後は、レスポンスの有効期間の判定である。この判定にはレスポンスの有効期限と生成されてからの経過時間が必要となる。まず、キャッシュは以下の値を有効期間として採用する。複数個存在する場合にはより上位に書かれている値を採用する。

- (1) キャッシュが共有キャッシュであり s-maxage が存在する場合、s-maxage の値
- (2) max-age が存在する場合、max-age の値
- (3) Expires ヘッダが存在する場合、(Expires の値 - Date ヘッダの値) で求められる値

もし、同項目に複数個の値が示されている場合、例えば、max-age に異なる値が2つ以上記述されている場合には、そのレスポンスは常に有効期限切れとして扱う。また、上記のどの値も得られなかった場合、キャッシュは任意の期限を設定する。この期限の設定では Last-Modified ヘッダの値を用いて計算することが多い。

次に、キャッシュが経過時間を計算するには以下の値が

必要となる。

$age$  = Age ヘッダの値

$date$  = Date ヘッダ値

$current$  = 現在時刻

$reqtime$  = リクエストの受信時刻

$restime$  = 格納レスポンスの受信時刻

これらの値を用いて、経過時間  $T_{age}$  は以下の手順で計算される。

$$T_{apparent} = \max(0, restime - date)$$

$$T_{corrected} = age + (restime - reqtime)$$

$$T_{initial} = \max(T_{apparent}, T_{corrected})$$

$$T_{visit} = current - restime$$

$$T_{age} = T_{initial} + T_{visit}$$

### 3.2 ヘッダのモデル定義

提案モデルでは、キャッシュおよび中継者の動作に必要なヘッダを新たに包括する。表??にその一覧を示す。

表 1 提案モデルで追加する HTTP ヘッダ

HTTP ヘッダフィールド	HTTP ヘッダ
HTTPRequestHeader	IfModifiedSinceHeader IfNoneMatchHeader PragmaHeader
HTTPResponseHeader	AgeHeader EtagHeader LastModifiedHeader
HTTPGeneralHeader	CacheControlHeader ConnectionHeader DateHeader WarningHeader
HTTPEntityHeader	ExpiresHeader

ヘッダフィールドはヘッダを用途と役割によって分類する。Ahkawe らによる既存モデル?では、HTTPRequestHeader と HTTPResponseHeader が含まれており、これらに加えて提案モデルでは HTTPGeneralHeader と HTTPEntityHeader ヘッダを追加している。この四種類の分類は HTTP の仕様においても用いられている。HTTPRequestHeader と HTTPResponseHeader はそれぞれ、リクエストもしくはレスポンスでのみ使用可能なヘッダが属する。これに対して、HTTPGeneralHeader と HTTPEntityHeader はリクエストとレスポンスのどちらにも使用可能なヘッダが属しており、この二つの違いは属するヘッダの役割である。HTTPGeneralHeader に含まれるヘッダはリクエストやレスポンスそのものの情報を表す。例えば、レスポンスのサーバにおける生成日時を記録するために用いられる。一方で、HTTPEntityHeader に属するヘッダはリクエスト

やレスポンスに含まれているコンテンツの情報を表す。例えば、レスポンスで送信されているファイルの最終更新日時をクライアントに伝達するといった役割を担う。

また、表??に含まれているヘッダのうち、特に重要な役割をもつ一部のヘッダを説明する。一つ目として Cache-Control ヘッダが考えられる。このヘッダは様々なオプションの記述によって、キャッシュの動作を指定することができる点で、キャッシュの動作に強く影響を与えるヘッダであるという点で重要である。使用するオプションの一例として max-age があり、非負整数を記述することでレスポンスがキャッシュに格納された際の有効期限を指定することができる。また、キャッシュへの保存を許可しない no-store、検証を行わないレスポンスの再利用を許可しない no-cache といったキャッシュの動作に制限を与えるオプションも存在する。

次に、キャッシュ内のレスポンスが再利用可能か判定するために必要なヘッダについて考える。再利用可能判定では、レスポンスの有効期限と生成からの経過時間が必要となり、経過時間の計算は??節において前述した計算式を用いる。したがって、この計算で用いる Date ヘッダ、Age ヘッダはレスポンスの再利用可能判定に必須であり重要なヘッダである。Date ヘッダはレスポンスのオリジンサーバにおける生成時刻、Age ヘッダはレスポンスの生成からの経過時間を表す。

最後にキャッシュの検証機能に必要なヘッダを考える。検証の動作はコンテンツの最終更新時間を用いる手法と、固有値を用いる手法の二種類が存在する。最終更新時間を用いる検証には、最終更新時間を記述する LastModified ヘッダと、記述されている時刻以降の更新が行われている場合にコンテンツの再送信を求める If-Modified-Since ヘッダが必須である。また、固有値を用いる検証には、固有値を伝達する Etag ヘッダと、記述されている固有値と最新のコンテンツの固有値が異なる場合に再送信を求める If-None-Match ヘッダが必要である。

提案モデル内には、役割が重複しているヘッダも存在している。例えば、Expires ヘッダはレスポンスの有効期限を、Pragma ヘッダは no-cache のオプションを表すことができるが、これらは Cache-Control ヘッダを使用することでも表現可能である。これは、Cache-Control ヘッダが HTTP/1.1 で新しく追加されたものであるからであり、HTTP/1.1 においては Expires ヘッダ、および、Pragma ヘッダの使用は推奨されていない。この二つのヘッダは、HTTP/1.1 が動作していない環境のみの使用が想定されていしかし、既存モデルは HTTP/1.0 と HTTP/1.1 の両方を含む環境を想定しているため、提案モデルにおいてもこの想定環境を引き継ぐことを目的としてこの二つのヘッダを包括している。

### 3.3 ウェブ要素のモデル定義

提案モデルでは、包括しているウェブの構成要素を図??のように構成している。図??において、濃く表示している要素は提案モデルで新たに追加した概念であり、矢印はクラスの継承構造を表す。

図??の一番上に存在する NetworkEndpoint はネットワークの末端に存在する要素の概念を表し、HTTPConfirmist は HTTP の仕様に従う動作をする概念を表す。既存モデルでは、クライアントとウェブサーバがこの概念に含まれており、この二つと並列に中継者の概念 (HTTPIntermediate) を追加している。さらに中継者はプロキシ (HTTPProxy) とゲートウェイ (HTTPGateway) の概念を内包している。

HTTP/1.1 においてプロキシは通過するコンテンツへの編集と通信経路の設定の二つの機能を持つが、これらの機能はキャッシュの機能に大きな影響は与えない。提案モデルはキャッシュの動作に注目した安全性検証を目的としているため、プロキシを通信経路上に存在するキャッシュを搭載する要素とみなす。

ゲートウェイは外部ネットワークから HTTP を用いてアクセスされる場合、ウェブサーバとして認識され同様の動作を行う。このことから、提案モデルではゲートウェイの内部ネットワークにおける動作は考慮せず、ウェブサーバとみなして設計する。

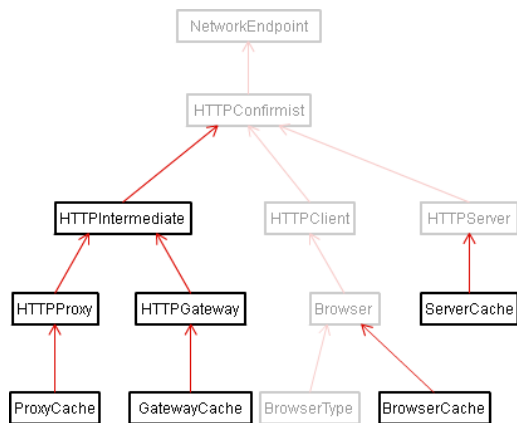


図 1 提案モデルに含まれるウェブアーキテクチャ

## 4. モデル実装

提案モデルの実装は形式手法ツールである Alloy を用いて進めている。Ahkawe らによる既存モデル?は Alloy を用いて実装が行われており、提案モデルはこれを拡張して実装を進める。

既に実装を終えている要素を図??にまとめる。これらの要素を用いることでキャッシュによるレスポンスの再利用や有効期限の判定を表現することが可能である。例えば、図??には CacheControl ヘッダが含まれている。CacheControl ヘッダの max-age のようないくつかのオプションには

整数値を用いるものが存在し、このようなオプションに整数値を表す Int クラスを関連付けている。同様に、日時を表す値を必要とする Date ヘッダにも Int を関連付けている。本来、日時を表す値と整数値は異なるものであるが、このように実装することは HTTP の仕様に反してはいない。なぜなら、日時を表す値を用いる際には、対応する整数値に変換して扱うためである。

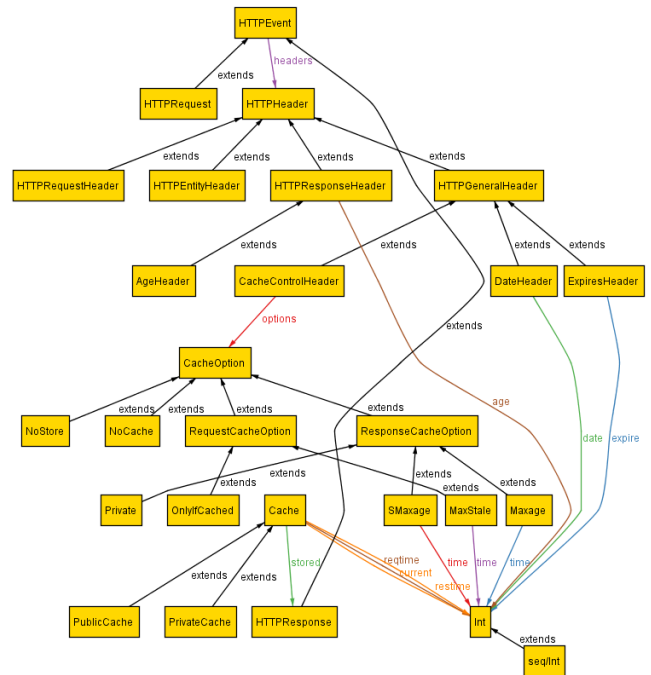


図 2 実装中の提案モデル

提案モデルの実装は未だ不完全であり、??章に前述した設計に従って実装を進めている。

## 5. 事例研究

### 5.1 レスポンスの格納と再利用

提案モデルのヘッダ定義が正当であることを確認する。図??は提案モデルの要件を満たす状態の一例である。

図??はある HTTP リクエストに対して、ブラウザキャッシュのような個別キャッシュに格納されているレスポンスを再利用し応答することが可能であることを示している。この HTTP リクエストのヘッダは以下のように想定されている。

Date: 6

同様に、HTTP レスポンスは以下のヘッダが想定されている。

Date: 1

Age: 7

Cache-Control: max-age=4

ここで、このレスポンスのキャッシュへの格納、およびこの状況における再利用が正当であるかを考察する。まず、このレスポンスにはキャッシュを禁止する記述はなく、か



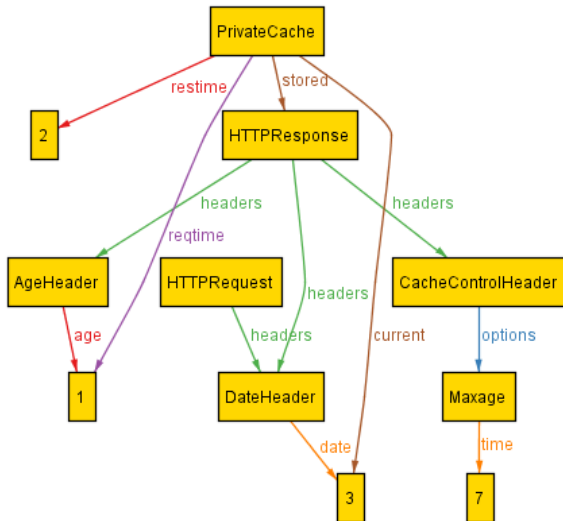


図 3 CacheControl ヘッダを用いたキャッシュの動作

つ、有効期限を明記する Cache-Control ヘッダの max-age オプションが指定されていたため、??節のキャッシュへの格納条件を満たしている。次に、??節のキャッシュの再利用条件について考察する。レスポンスのヘッダに有効期限を過ぎているレスポンスの再利用を許可する記述はなく、キャッシュの検証も行っていないため、レスポンスが有効期限内である必要があり図??の情報から判定を行う。有効期限は max-age の値が採用され 4 となり、レスポンスの経過時間を??節の計算式を用いて求める。

$$age = 7$$

$$date = 1$$

$$current = 5$$

$$reqtime = 2$$

$$restime = 1$$

$$T_{apparent} = \max(0, restime - date)$$

$$= \max(0, 0) = 0$$

$$T_{corrected} = age + (restime - reqtime)$$

$$= 7 + (1 - 2) = 6$$

$$T_{initial} = \max(T_{apparent}, T_{corrected})$$

$$= \max(0, 6) = 6$$

$$T_{visit} = current - restime$$

$$= 5 - 1 = 4$$

$$T_{age} = T_{initial} + T_{visit}$$

$$= 6 + 4 = 10$$

## 5.2 BCP 攻撃

## 6. 結論

謝辞 謝辞謝辞謝辞