コンピュータセキュリティ 11－4
（2000．9．8）

プライバシ保護を考慮した地理位置情報システム

渡辺 恭人 [1] 慶應義塾大学政策・メディア研究科
竹内 奏吾 [2] (株) ソニーコンピュータサイエンス研究所
寺岡 文男 [3] (株) ソニーコンピュータサイエンス研究所
村井 純 [4] 慶應義塾大学環境情報学部

概要

プライバシー保護を実現した地理位置情報管理システム（GLI システム）を提案する．我々が提案している GLI システムは，インターネットに接続している移動体の地理位置情報を地球規模で管理するものである．移動体は自分の位置をサーバに登録し，検索者は移動体の識別子を鍵とした位置検索，および地理範囲を鍵とした移動体検索が可能である．GLI システムは地球規模での動作を可能とするため，サーバの階層化による分散管理を導入している．本稿で扱うプライバシー保護とは，第三者による移動体の特定防止，位置特定防止，追跡防止，なりすまし防止およびインターネットにおける盗聴防止，データ改竄防止である．さらにサーバが管理するデータベースの盗難も考慮する．移動体の識別子として HID(hashed ID) を導入することにより，信頼関係のある検索者のみに移動体の特定，位置特定，追跡を可能とし，信頼関係のない検索者には統計情報のみの利用を可能とする．なりすまし防止のため位置登録サーバを導入し，移動体の認証を行う．インターネットにおける盗聴防止および改竄防止には IPsec を利用する．

The Geographical Location Information System with Privacy Consideration

Yasuhito Watanabe Keio University
Sohgo Takeuchi Sony Computer Science Laboratory
Fumio Teraoka Sony Computer Science Laboratory
Jun Murai Keio University

Abstract

We propose the Geographical Location Information (GLI) System with Privacy consideration. The GLI System we propose provides a way to manage geographical location information of mobile entities in the worldwide scale. Mobile entities regiter their location information with servers. Searching clients are able to look up location of mobile entities using specified identifier as a search key and look up identifiers of mobile entities using speified geographical region as a search key. The GLI System has a distributed management method in order to be scaled to the world.

In this paper, we introduce HID (Hashed ID) to the GLI System for privacy preservation. The privacy preservation means that a mobile entity must not be identified by unknown persons, the location of a mobile entity must not be realized, and it is impossible to track a mobile entity. We designed the new architecture of GLI System named GLIsec.

[1] riho-m@sfc.wide.ad.jp
[2] sohgo@csl.sony.co.jp
[3] tera@csl.sony.co.jp
[4] jun@sfc.wide.ad.jp

1 はじめに

我々が提案している地理的位置情報システム（GLI: Geographical Location Information System）[1][2]では，現実世界を移動する移動体を対象とし，その識別子と位置情報の登録・検索機能の実現を目指している．移動体は，その位置情報，付帯する状態や属性に関する情報を持つ．本システムにより，計算機やユーザの位置・状態をインターネットを通じて認識することができる．移動体はサーバに位置や状態の情報を登録し，クライアントは，識別子や位置を鍵とした検索要求をサーバに送信することにより，移動体を検索することができる．文献[2]ではサーバの階層化による分散管理を実現した．

本稿ではさらにプライバシ保護機能をGLIシステムに付加する，プライバシの保護とは，移動体が登録した情報を第三者から隠蔽することを意味する．単に識別子を隠蔽しただけでは，どの検索者も移動体を特定できなくなる．本システムの検索の一つである地理的位置に基づいた検索は，地理的位置情報が隠蔽されている場合は行えなくなるため，識別子のみの隠蔽が必要である．識別子を隠蔽する場合には，不特定多数の第三者からは隠蔽し，信頼する特定少数に対しては隠蔽しないといったアクセスの制御が必要となる．これらにより，既存の地理位置情報システムの利点である集まった分布情報に対する検索機能を損わずに可能となる．

本論文では，これまでの地理位置情報システムの利点である検索手法を損なうことなく，プライバシー保護を実現した地理位置情報システムを提案する．

2 地理位置情報システムの概要

地理位置情報管理システム（GLIシステム）は，インターネットに接続している移動体の識別子と地理位置情報および付帯情報の管理・検索機能を提供する．移動体の識別子としてはFQDN（Fully Qualified Domain Name）[1]を用いる．地理位置情報としては緯度・経度・高度を用いる．付帯情報とは，移動体の移動方向や移動速度などである．GLIシステムのサーバ（GLIサーバ）は，移動体が登録した地理位置情報や付帯情報を管理する．またGLIサーバは2種類の検索機能を提供する．1つは移動体の識別子を鍵とし，その移動体の位置情報および付帯情報を返すものである（正引き検索）．もう1つは地理的な領域を指定し，その領域に存在する移動体の識別子，位置情報および付帯情報の集合を返すものである（逆引き検索）．

図1にGLIシステムの構成と動作例を示す．GLIサーバは，正引き機能を提供するホームサーバ群と，逆引き機能を提供するエリアサーバ群からなる．それぞれのサーバ群は階層構造をとっており，分散管理[2]によって大規模性を実現している．

[1]ドットで区切られたホスト名．例えば，mobile.sfc.wide.ad.jp

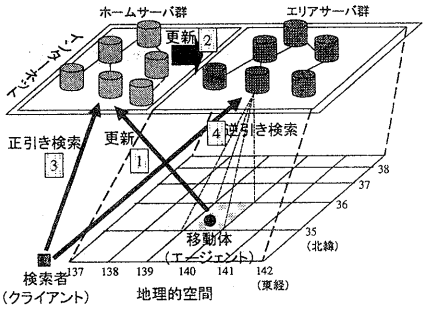

図1: GLIシステムの構成

図ではmobile.sfc.wide.ad.jpという識別子を持つ移動体が北緯35度18分18秒，東経139度30分40秒に存在している．移動体は識別子から決定されるホームサーバに識別子，地理位置情報および付帯情報を登録する（図1-（1））．登録を受けたホームサーバは，移動体の地理位置情報から決定されるエリアサーバに，移動体の識別子，地理位置情報および付帯情報を登録する（図1-（2））．

正引き検索を行う検索者は，移動体の識別子から決定されるホームサーバにmobile.sfc.wide.ad.jpという識別子を鍵として検索要求を送信する（図1-（3））．検索要求を受信したホームサーバは，検索結果として北緯35度18分18秒，東経139度30分40秒という地理位置情報および付帯情報を返す．逆引き検索を行う場合は，例えば，北緯35度～36度，東経139度～40度という領域を鍵とし，この領域から決定されるエリアサーバに検索要求を送信する（図1-（4））．検索要求を受信したエリアサーバは，mobile.sfc.wide.ad.jpという識別子，北緯35度18分18秒，東経139度30分40秒という地理位置情報および付帯情報を返す．指定された領域に他の移動体も登録されている場合は，その情報も返す．

3 GLIシステムでのプライバシ保護

前提条件

図2に地理位置情報がおかれる環境とその脅威を示す．GLIサーバ群と移動体および検索者はインターネットを介して通信する．GLIシステムでのプライバシ保護を考える上で，以下の2つの前提を置く．1つは，インターネットでの通信では基本的にセキュリティが考慮されていないことである．もう1つはGLIサーバは不正を行わないということである．また，移

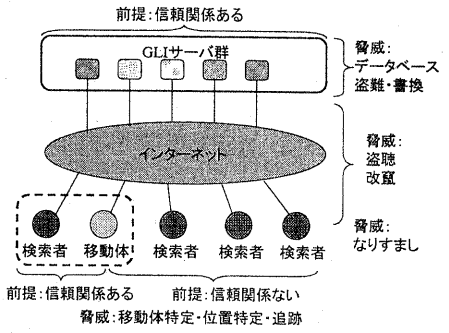

図 2: GLI システムにおける環境と脅威

動体にとって，信頼関係にある検索者と信頼関係のない検索者（第三者）という区別があるものとする．このような前提条件のもと，GLI システムにおけるプライバシ保護の目標を以下に述べる．

移動体の特定

GLI システムにおける移動体のプライバシとは，移動体が特定されその位置情報が明らかになることである．従って，元々の移動体の識別子とは関連のない識別子を別に用意するなどの方法によって識別子の隠蔽を行い，自分以外の第三者には，移動体と位置情報の対応関係を隠蔽する必要がある．つまり，移動体の位置情報と付帯する情報は公開するが，移動体の特定はできないことを目標とする．

しかし，第三者から移動体の識別子を隠蔽するだけでは，自分以外の全てからプライバシを保護することができるが．実用的ではない．ある移動体と信頼関係にある検索者が，その移動体を検索し位置を知ることにより，応用可能性が増す．このように，信頼関係にある検索者からのみ移動体を特定しての検索を可能とし，それ以外の第三者はある領域に存在する移動体の集合に関する統計情報のみを取得できる．

移動体の追跡

逆引き検索により，検索者はある移動体についてその特定はできないが，位置を追跡することは可能となる．追跡が可能となれば，移動体の移動範囲からその移動体が特定されることも考えられる．したがって，移動体の追跡防止も目的とする．

なりすましによる偽の移動体情報登録

悪意のある移動体が別の移動体になりすまして GLI サーバに偽の位置情報を登録することも考えられる．GLI システムでは位置情報を登録する移動体を認証し，正しい情報のみを扱うことを目的とする．

データベース盗難

インターネットのようないくつものネットワークが相互に複雑に接続されているネットワークで，ユーザも多数存在する環境では，さまざまな攻撃による乗っ取りからデータベースの盗難といった事故の可能性がある．システム内部では，移動体と位置情報の対応関係を隠蔽し，データが盗まれても意味をなさないことを目標とする．

盗聴・改竄

移動体はネットワークを介して GLI サーバに自らの位置情報を登録する．ネットワークを介した通信を行うことで，通信データを盗聴されたり改竄される可能性がある．やりとりされるパケットには移動体の送信元のアドレスが含まれるため，データ部分に含まれる登録情報との対応関係が明らかになる．データ部分を隠蔽するなどの方法により，盗聴されても移動体と位置情報との対応関係を明らかにさせないこと，また，通信データが改竄された場合は GLI サーバがその事実を検出できるようにすることを目標とする．

4 解決手法

Hashed ID の導入

移動体と検索者間における信頼関係の有無によって，特定可能性・不可能性を制御するには，移動体と信頼関係にある検索者だけが理解できる秘密の識別子を移動体と共有し，この秘密の識別子を GLI サーバに登録すればよい．しかしこれだけでは，第三者が逆引き検索から得た秘密の識別子を使って正引き検索をすることにより，追跡が可能となる．追跡を防ぐには，秘密の識別子を頻繁に変更する必要があるが，変更のたびに移動体と検索者の間での通信が必要となり，通信によるオーバヘッドが大きくなる．したがって，移動体と検索者が通信することなく，頻繁に変化する秘密の識別子を共有する方法が必要となる．そこで両者の時刻同期を前提として，鍵付きハッシュ関数の鍵として時刻情報を使用する方法を提案する．ハッシュの結果得られる値を HID（Hashed Identifier）とし，これを秘密の識別子として利用する．時刻情報としては，基準時刻（ts）と基準時刻変更の間隔（ttl）を導入する．ハッシュ関数を利用することにより得られた出力値から元の識別子へ復元することは不可能である．

もともとの識別子を ID とすると HID は次のような式で表される．

$$HID = hash(ID \oplus (ts + ttl * n)) \quad (n \text{ は } 0 \text{ 以上の整数})$$

n の値は現在時刻から計算でき，$ts + ttl * n <$ 現在時刻 $< ts + ttl * (n+1)$ のような関係となる．この式から求められる HID の値が時刻によって変化する過程を図 3に示す．移動体は，信頼関係にある検索者と HID 生

成に必要な情報 ID, ts, ttl を安全な通信路を使用して事前に共有しているものとする．最初の登録 $(T = ts)$ 時に，移動体は $HID0 = hash(ID + ts)$ と位置情報等をサーバに送信する．$ts + ttl <= T < ts + 2 * ttl$ では，$HID1 = hash(ID + ts + ttl)$ となる．このように，HID は時間とともに変化する．ある移動体と信頼関係にある検索者は，その移動体と時刻が同期しているので，同時刻に同じ HID を生成することができ，検索の鍵として使用する．

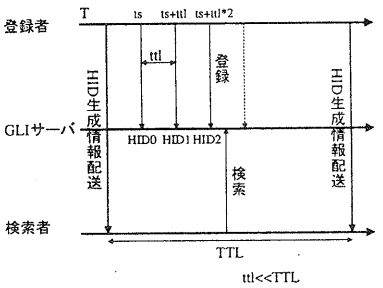

図 3: HID 生成の時間変化

しかし ttl だけの場合，識別子と ts および ttl が何らかの理由によって第三者に洩れた場合，第三者は時間とともに変化する HID の生成が可能となってしまう．したがってさらに ttl より大きい一定間隔（TTL）で信頼関係にある移動体と検索者間で鍵の元になる情報を新たに交換し，鍵を変更することにより，安全性が保たれる．

HID を用いた登録と検索

HID を導入した GLI システムにおける登録・検索に関して述べる。まず，正引き検索，つまり HID からの検索について述べる．

登録者となる移動体 A と，A と信頼関係にある検索者 B．この 2 者では，HID 生成に必要な情報を共有する．A と信頼関係にない検索者 C は，A の HID 生成に必要な情報は持たないとする．まず A は識別子 ID_a からハッシュ関数を利用して A の HID として，HIDa を得る．A は HIDa と A の位置情報を GLI サーバに送信する．サーバには HIDa と位置情報が登録される．次に検索者 B は，A の位置情報を知りたいので，A の識別子 ID_a からハッシュ関数により HIDa を生成し，GLI サーバに対して，HIDa を検索の鍵として検索要求を行う．サーバは登録されている HID から HIDa にマッチするものを検索し，その位置情報を検索者 B に送信する．検索者 C は A の HID として HIDa を生成できないため，検索要求することができない．

次に逆引き検索について述べる．逆引き検索はある位置領域を指定し，その領域に属する移動体を検索する．登場する移動体と検索者は正引き検索と同様である．GLI サーバにはすでに A の HID と位置情報が登録されているとする．A と信頼関係にある検索者 B がある位置領域を指定してその領域に属する移動体を検索要求する．サーバでは該当する移動体の HID とその位置情報をリストとして B に送信する．B では，検索結果の HID を B で計算できる HID と比較し，HIDa を発見し，A がその領域に属していたことがわかる．C は同様の検索を行い，結果として同じリストを受信する．しかし，C は HIDa を計算できないので，受信した HID に対応する移動体の特定はできないが，公開情報である位置情報，その領域に属する移動体の数などの統計情報を得ることができる．

HID 生成のハッシュ関数

HID の生成に用いる一方向性ハッシュ関数は，これら生成に必要な追加の情報を鍵として利用することを考えると，鍵付きハッシュ関数（HMAC: Keyed Hashing for Message Authentication）[4] が適当である．また HMAC と組み合わせて使用するハッシュ関数としては，処理速度やセキュリティの強度から SHA-1（Secure Hash Algorithm）[3] が適当である．SHA-1 は現状では異なる入力値から同一の出力値を作る可能性がほとんどなく，出力値に偏りもないハッシュ関数で MD4 や MD5 に比べて一意性の点で優れている．したがって，HID の生成には，HMAC-SHA1[4] を使用する．

HMAC-SHA1 は，入力値として任意の文字列，160bit の鍵により，160bit の値を出力する．本提案では，入力値に ID，鍵には ts（32bit）と ttl（32bit）の使用を想定しているが，時刻情報以外に予測不可能な情報を加えることはセキュリティの強度を高めるという点で重要である．予測不可能な値として残りの 96bit は乱数を鍵として使用する．

移動体の認証機能の導入

移動体が登録を行う際にはその移動体を認証する必要がある．そのため，移動体を認証し，移動体の情報は登録せずにデータベースを持つ HID サーバに転送を行う登録サーバを導入する．データベースを持つ HID サーバは登録 HID によって固定されないため，認証が困難になる．登録サーバを導入することにより，移動体は常に同じ登録サーバと認証し，移動体と移動体の情報を登録する HID サーバとの関連を無くすことができる．また移動体と登録サーバの通信が盗聴された場合，送信元である移動体のアドレスと HID が対応付けられる．このような認証と機密性を確保するには，IPsec[5] の機能を利用することで解決できる．IPsec には AH（Authenticatin Header）と ESP（Encapsulating Security Payload）の二つのプロトコ

ルからなる．AH はパケットのヘッダにハッシュ署名を使用することによって，パケットデータの完全性と送信者認証を証明するプロトコルである．ESP は，暗号機構を使用して完全性と発信元認証と機密性を確保する認証・暗号化プロトコルである．移動体を認証し登録サーバとの通信の機密性，完全性を確保するために ESP を使用する．

データベースの配置

これまでの GLI システムと同様に，正引きの検索，逆引きの検索，それぞれの検索を受け持つサーバが必要である．本論文で提案するシステムでは，正引きは HID を鍵とした移動体の検索であり，逆引きは位置情報を鍵とした移動体の検索となる．それぞれのサーバでは，移動体の情報をデータベースに蓄積するが，データベースが盗難にあった場合に，盗まれた情報から移動体が特定できてはならない．またそれぞれのデータベースには HID と位置情報以外の移動体に関連する情報は置かないことで，HID の持ち主の特定を不可能にする．また登録サーバでは管理している移動体の情報のうち，HID と位置情報以外の情報を持つことにより移動体を特定することを防ぐ．このように，保持されるデータを分離して関連性をなくすことにより盗難に対する安全性を高める．

盗聴・改竄防止機能の導入

サーバ間の通信，検索での通信において，盗聴・改竄を防ぐには．IPsec の ESP を利用することで解決できる．サーバ間の通信でやりとりされるのは HID と位置情報と ttl だけであり，これらの情報から移動体は特定されないため盗聴されてもよい．サーバ間では，IPsec における AH を利用して改竄を防止する．

5 GLIsec システム

前章までに述べた解決手法から，これまでの GLI システムを拡張し，図 4のような GLIsec システム構成を導入する．

本システム上で管理する情報を挙げる．GLI（Geographical Location Information）は移動体の地理位置情報であり，緯度，経度，高度によって指定される 1 地点を示す．GLI は公開情報として扱う．ID（identifier）は移動体の識別子であり，FQDN や IP アドレスなど，形式は任意である．HID（hashed identifier）は，ID をもとにハッシュ関数（HMAC-SHA1）によって生成される数値で 160bit の長さを持つ．移動体はこれを GLI とともにサーバに送信し登録する．検索者は信頼関係にある移動体の HID を生成でき，正引き検索に使用する．

本提案システムは，登録サーバ群，HID サーバ群，エリアサーバ群という 3 種のサーバ群と登録クライアント，検索クライアントという 2 種のクライアントから構成される．

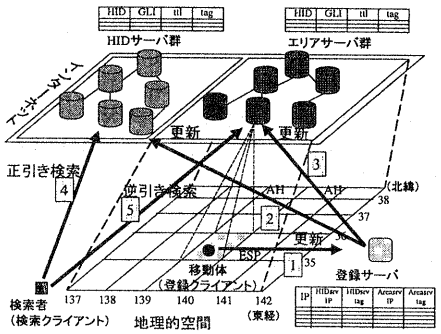

図 4: GLIsec システム構成

登録クライアントは，移動体で動作するプログラムで，GPS などの位置取得装置から GLI を取得し，識別子と鍵から HID を生成する．GLI と HID と有効期限である ttl を特定の登録サーバへと送信する（図 4-(1)）．

登録サーバは，登録クライアントからの GLI 登録を受け付けるプログラムである．登録サーバは，GLI 登録を受け付ける登録クライアントを認証し，特定の登録クライアント以外からの登録は受け付けない．登録サーバは，受け付けた位置情報登録を蓄積せず，HID の値，GLI の値に従って特定される HID サーバ，エリアサーバに対して，HID，GLI，ttl を送信する（図 4-（2）（3））．

登録サーバは，登録クライアントから登録を受け付ける際にその IP アドレス，HID の値によって特定される HID サーバの IP アドレス，GLI の値によって特定されるエリアサーバの IP アドレスを保持する．また，HID サーバ，エリアサーバへの登録時に返されるデータベースエントリの tag も保持する．

HID サーバは，登録サーバから HID，GLI，ttl を受信して蓄積する．また，検索クライアントからの HID を鍵とした移動体の検索，正引き検索を受け持つ．既存の GLI システムにおけるホームサーバのデータベース部の役割を持ち，正引き検索を受け持つサーバである．HID サーバは，HID の値によって登録する HID サーバが決定され，移動体にとって HID サーバは固定されない．HID サーバのは HID による階層化により分散管理する．

エリアサーバは，サーバは特定の位置領域に存在する移動体の HID，GLI，ttl を登録サーバから受信して蓄積する．また，検索クライアントからの位置領域を鍵とした移動体の検索，逆引き検索を受け持つ．エリアサーバ群は受け持つ領域に従って木構造をなし，その管理構造は緯度・経度によって分割されたメッシュ

によって構成される木構造を利用した分散管理を行う[2].

検索クライアントは，正引き・逆引きの検索要求をサーバに対して行いその結果を受信するプログラムである（図4-（4）（5））．ある信頼関係にある移動体とHIDを生成するための情報を共有している検索者と，そうでない検索者が存在する．前者は移動体をHIDにより特定できるため特定可能検索クライアントとし，そうでない検索者，つまりある移動体のHIDを生成するための情報を持たない検索者が使用する検索クライアントを特定不能クライアントとする．

6 考察

時刻同期の妥当性と登録の頻度（ttl）

登録クライアントと特定可能検索クライアントでは，時刻情報によりHIDを同期して生成する．両者では時刻を同期する必要がある．時刻同期の手法としては，ネットワーク経由で時刻同期を行うNTPがある．NTPを使用してインターネット上で時刻を伝送する場合の誤差は20ms程度，ネットワークの混雑により増加する可能性はあるが，現状の時刻同期手法としては最も有用である．ttlの値は1分から5分程度の時間を想定しており，NTPの誤差と比較して大きい値である．したがって誤差によってHIDを同期できない可能性は低い．

HIDの衝突

HIDはHMAC-SHA1という鍵付きハッシュ関数を利用して生成する．HMAC-SHA1による出力値の衝突可能性は，使用されるハッシュ関数SHA-1の衝突可能性に依存する．文献[3]によると，異なる入力値に対して、同一のハッシュ値を返す可能性がほとんどないことがわかっている．登録しようとしているHIDがすでに使用されている場合はエラーとして登録クライアントに通知し，識別子や鍵を変更して再度登録を試みる．

登録サーバにおける盗聴

本提案のシステムでは，登録サーバを導入することにより，登録クライアントの認証を行い，データベースであるHIDサーバと登録クライアントの関連を無くしている．登録サーバが動作するホストが接続されるネットワークにおいてパケットの盗聴が行われた場合を考える．登録クライアントから登録サーバへの通信はIPsecのESPにより暗号化されるが送信元のアドレスは明らかになる．登録サーバからHIDサーバ，エリアサーバへの通信はIPsecのAHを使用するため，パケットの内容（HID, GLI, ttl）は明らかになる．従って，登録クライアントのアドレスとHIDは対応付けられる可能性はある．

各サーバにおけるデータベース

登録サーバには，管理している移動体のアドレス，登録しているHIDサーバ，エリアサーバの情報を保持している．HIDサーバ，エリアサーバでは，HIDと位置情報を蓄積している．これらのうち，一つのデータベースが盗まれた場合は，移動体を特定することはできない．登録サーバを含む二つのデータベースが盗まれた場合はその対応付けから移動体を特定できる．

7 まとめと今後の課題

本論文では，移動体のプライバシ保護を考慮したGLIシステムを提案した．本提案方式では，従来ともに公開情報となっていた移動体の識別子と位置情報のうち，識別子をハッシュ関数に入力させることで得られた値をHID（Hashed ID）として利用する．これにより第三者から識別子を隠蔽し特定を不可能にした．また，ハッシュ関数の入力値として時間変化する鍵を信頼関係のある検索者と共有することにより，信頼関係のある検索者には特定を許し，第三者からの追跡可能性を減少させることができた．

本論文では，提案方式の設計について述べた．今後はこの設計を基にした実装を行い，大規模運用可能な地理位置情報管理システム[2]との協調による実験運用を行った上での性能評価，さらに，HIDサーバ群の分散管理構造，鍵の変更と配送などに関する検討を行う．

参考文献

[1] Yasuhito Watanabe, Atsushi Shionozaki, Fumio Teraoka, Jun Murai, “The design and implementation of the geographical location information system.”, Proc. of INET'96. Internet Society, June 1996.

[2] 竹内 奏吾, 中村 嘉志, 多田 好克: インターネットにおける地理位置情報管理システムの設計と実装, 情報処理学会 マルチメディア，分散，協調とモーバイル（DICOMO'99）シンポジウム論文集, pp. 405-410, June, 1999

[3] National Institute of Standards and Technology （NIST）, FIPS PUB 180-1: Secure Hash Standard, April 1995.

[4] Krawczyk, H., Bellare, M. and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, February 1997.

[5] Kent, S., and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998