社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2008－IOT－1（6）
2008/5/8

IPマルチキャストを用いた放送型暗号による ライブ映像配信システムの構築と評価

森村 吉貴† 上原哲太郎†† 侯 書会† 美濃 導彦††

† 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町
†† 京都大学学術情報メディアセンター 〒606-8501 京都市左京区吉田二本松町
E-mail: †{morimura,shuhui}@mm.media.kyoto-u.ac.jp, ††{uehara,minoh}@media.kyoto-u.ac.jp

あらまし 我々はIPマルチキャスト上のSRTPを用いた鍵配送に放送型暗号を用いたライブ映像配信システムを構築した．一般に認証型のライブ映像配信では，ユーザ鍵の流出と復号済映像の流出という問題が残る．我々のシステムは，ユーザ鍵の流出については放送型暗号のユーザ鍵が異なるという特徴を生かし，復号済映像については透かしとして映像にIDを埋め込む電子指紋により，それぞれの流出を抑止する．実験では構築したシステム上で帯域資源と計算資源のオーバーヘッドを計測し，ライブ映像配信に十分に適用可能であることを示す．
キーワード ライブ映像配信，著作権保護，放送型暗号，電子指紋

Construction and Evaluation of Live Video Streaming System with Broadcast Encryption over IP Multicast

Yoshitaka MORIMURA†, Tetsutaro UEHARA††, Shuhui HOU†, and Michihiko MINOH††

† Graduate School of Informatics, Kyoto University
Yoshidahonmachi, Sakyo-ku, Kyouto-shi,Kyoto, 606-8501, Japan
†† Academic Center for Computing and Media Studies, Kyoto University
Yoshidanihonmatsu-cho, Sakyo-ku, Kyoto-shi, Kyoto, 606-8501, Japan
E-mail: †{morimura,shuhui}@mm.media.kyoto-u.ac.jp, ††{uehara,minoh}@media.kyoto-u.ac.jp

Abstract We constructed a live video streaming system over IP Multicast, which uses SRTP and delivers the decryption key with broadcast encryption. Generally, live video streaming with authentification has problems as illegal distribution of user keys and decrypted video. Our system deters user key distribution using characteristics of broadcast encrpytion, which has different user keys by users, and deters decrypted video distribution using digital fingerprint, which embeds user ID to video as watermark. We evaluated overhead of computational resource and bandwidth resource on the constructed system and show the system works in realtime enough for live video streaming.
Key words Live Video Streaming, Copyright Protection, Broadcast Encryption, Digital Fingerprinting

1. はじめに

今世紀に入り，インターネットにおける映像配信はYoutubeに代表されるように急速な普及を見せた．電波放送やケーブルテレビなどの既存の映像配信設備を持たない集団や個人でも，CGM(Consumer Generated Media)として映像による創作活動を活発に行うようになっている．この流れに応じ，地域イベントやサークル活動，教育機関による遠隔教育などを，ネットワークを通じてライブ配信することへのニーズが高まっている．映像は撮影された被写体に関し，著作権や肖像権などの権利を多く含んでいるため，著作権保護は重要な課題である．映像の有料配信やコミュニティ内に閉じた配信を行いたい場合，正しく認証されたユーザのみが閲覧可能なライブ映像放送システムの構築が必要である．本研究では，既存の映像放送設備が対象としないような，数百人程度から数万人程度までのユーザ数を対象とする，同期的な認証型ライブ映像配信システムの構築を目標とする．

認証型のライブ映像配信において，インターネットのようなオープンな経路を利用する際には，配信側でコンテンツを暗号化し，正しく認証されたユーザに復号鍵を配布する(以降，ユー

ザに配布された復号鍵を**ユーザ鍵**と呼ぶ) ことによって，配信経路上の流出という問題は防止することができる．しかし，正しく認証されたユーザが，ユーザ鍵や復号済み映像を不法に再配布し，認証されていないユーザからも映像を閲覧可能にする問題は残っている．この行為を海賊行為といい，Youtube や P2P などを利用した海賊行為は既に巨大な規模となっていることが判明しており，ライブ配信においても同様の行為が行われることが懸念されている．

このような海賊行為を抑止するためには，ユーザ鍵や復号済映像を誰が流出させたかを追跡可能することが有効であると考えられる．流出者が追跡可能であれば，その法律的・社会的な責任が問えるようになり，流出者にとって海賊行為が見合わなくなると考えられる．流出者を特定可能とするためには，ユーザ鍵や復号済映像がそれぞれユーザごとに異なる ID を持つようにし，それらが流出した際は発信者の持つ ID 情報との照合により流出者が特定できるようにすればよい．次節では，ライブ映像配信において多数のユーザに対し，異なるユーザ鍵の配布と映像への ID 情報の挿入を行う手段を述べる．

2. 認証型ライブ映像配信におけるユーザ鍵と復号済映像の流出抑止と関連研究

まず，ライブ映像配信の枠組みとして，一人の発信者が，ネットワークを介して多数のユーザに対し映像を発信する状況を想定する．多数のユーザに対してライブ映像配信を行うため，計算資源と帯域資源を最低限に抑えリアルタイム性を保持する必要がある．流出防止を目的として多数のユーザに異なるユーザ鍵を配布するための方法として，まず，暗号化に利用する鍵 (以降，これを**サーバ鍵**と呼ぶ) とユーザ鍵のペアをユーザ数と同数だけ準備し，ユーザごとに異なる暗号化ストリームを配信することが考えられる．この方法の問題は，多数のストリームの暗号化・配信が必要となり計算資源・帯域資源を著しく圧迫することである．この問題は，映像発信者が持つ単一のサーバ鍵で暗号化したストリームを複数の異なる鍵によって復号することができる性質を持つ暗号方式を用い，単一の暗号化ストリームを IP マルチキャストによって暗号化することにより解決できる．Fiat らによって提案された放送型暗号 [1] はそのような性質を持っており，提案するシステムでは放送型暗号を利用する．一方，復号済の映像の流出を防止するためには，映像信号自身に ID 情報を埋め込む必要がある．このような利用者の識別を目的としてコンテンツ自身に埋め込まれた ID 情報は，電子指紋と呼ばれる．対象が画像の場合，画質に影響なく電子指紋を挿入するために，電子透かしが用いられる．今回提案するシステムでは対象が映像であるため，高いリアルタイム性を持つ電子透かしが必要とされる．

そこで本研究では以下のように，IP マルチキャストを用い，放送型暗号による暗号化と電子指紋の挿入の機能を備えたライブ映像配信システムを提案する．発信システムはサーバ鍵とユーザ鍵の組を生成し，映像配信を行う前にあらかじめユーザにユーザ鍵を配布しておく．ライブ配信開始時には発信システムは放送型暗号により暗号化を行ったストリームと復号に必要

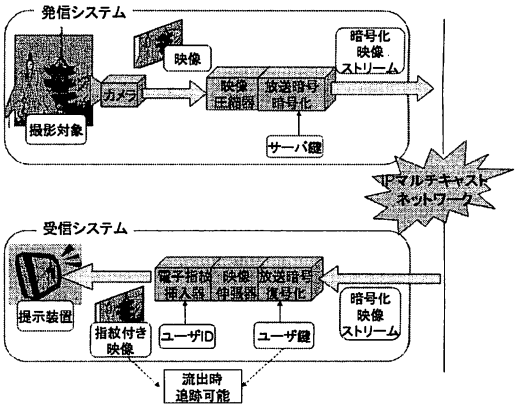

図 1　提案システムの概要図

なヘッダ情報をネットワークへ IP マルチキャスト送信し，ユーザ側の受信システムはネットワークからマルチキャストにより受信したストリームをヘッダ情報と自身のユーザ鍵とによって放送型暗号により復号し，各人の ID を電子指紋を挿入する．この手続きを図 1 に示す．

放送型暗号の実装の例として，Fiat-Naor らによる方式の実装 [3] や Boneh らによる方式 [2] があるが，実際の映像放送に対し適用した際の具体的なユーザ数を想定した計算量の評価と実装はなされていない．また，PC 上でリアルタイムで映像にソフトウェアで透しを埋め込む方式の研究も提案されている [4] が，発信側での電子透しの埋め込みを想定して映像圧縮とは異なる専用の PC の利用を想定したものであり，本研究のように，受信側で閲覧時に放送型暗号と併用されながら挿入する用途を対象としたものではない．本研究では放送型暗号による暗号化と電子指紋の挿入を共に行う IP マルチキャ ストライブ映像配信システムを PC 上でソフトウェア的に構築し，全ての処理をリアルタイムに行うシステムを試作した上で帯域資源と計算資源の利用に関する評価を行った．

3. 放送型暗号による暗号化と電子指紋の挿入を行う IP マルチキャストライブ映像配信システムの構築

放送型暗号システムには，結託攻撃に強い耐性を持つ Boneh, Gentry, Waters らによる放送型暗号 [2] を用いることとする．以降，Boneh らの方式を BGW 放送型暗号と呼ぶ．BGW 放送型暗号を用いれば，単一のサーバ鍵から特定の符号列とそれを暗号化したものを生成することができる．ここではこの処理を BGW 暗号化と呼ぶ．暗号化された符号列はユーザ鍵を用いることで復号することが可能で，この処理を BGW 復号化と呼ぶ．このとき，特定の符号列は送信者が選択することができないため BGW 暗号化自体は任意の映像ストリームを暗号化することはできない．しかし，BGW 暗号化した特定の文字列をネットワークを通じて配布し受信側でそれを BGW 復号化し，発信システムと受信システムがその特定の文字列を共有して既存の共通鍵暗号システムの鍵として用いることで，間接的に任

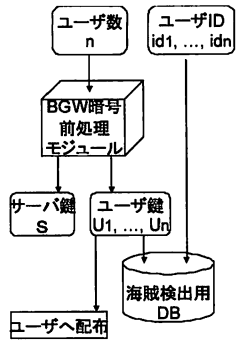

図 2　前処理システム

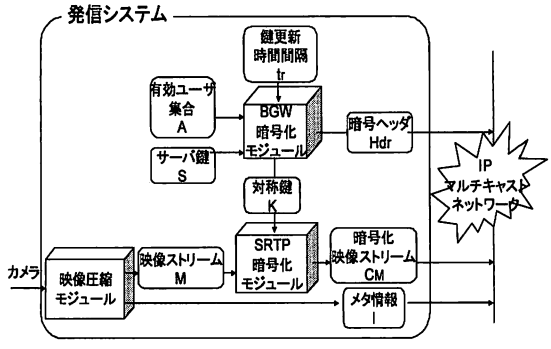

図 3　発信システム

意の映像ストリームを暗号化・復号化することが可能である．以降この特定の文字列を**対称鍵**，それを BGW 暗号化したものを**暗号ヘッダ**と呼ぶ．

共通鍵暗号によって暗号化された映像ストリームをリアルタイム配信するのに適したプロトコルに，RFC3711 で規定された SRTP [5] がある．提案するシステムは，SRTP プロトコルを基盤とし，BGW 暗号方式によって生成された対称鍵を SRTP 共通鍵方式の鍵として用いる IP マルチキャスト配信を行う．以下では，提案するシステムの前処理，発信システム，受信システムについて詳細に述べる．

提案するシステムにおける発信側の前処理を図 2 に示す．発信者は，発信前に前処理としてユーザ数 n を指定し，BGW 暗号前処理モジュールによってサーバ鍵 S 及びユーザ鍵 $\{U_i|i=1,\ldots,n\}$ を生成し，ユーザ鍵 U_i を認証されたユーザ i に配布する．このとき，発信者はユーザ ID id_i とユーザ鍵 U_i を海賊検出用データベースに記録し，認証されたユーザ i と紐付けておくものとする．

発信システムは映像圧縮モジュール，BGW 暗号化モジュール，SRTP 暗号化モジュールからなる (図 3)．映像圧縮モジュールは PC に取りこまれた映像の圧縮符号化を行い映像ストリーム M として出力する．発信システムは有効ユーザ集合 $A\subseteq\{1,\ldots,n\}$ を指定し，BGW 暗号化モジュールによってサーバ鍵 S を元に対称鍵 K と暗号ヘッダ Hdr を生成する．次に，SRTP 暗号化モジュールが対称鍵 K を用いて映像ストリーム M を暗号化

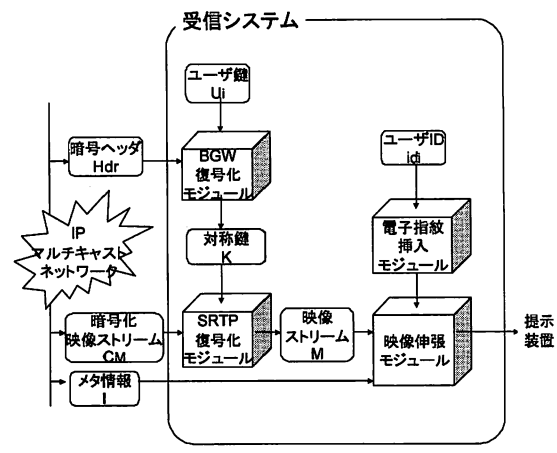

図 4　受信システム

した上で RTP 形式に整形した暗号化ストリーム C_M をネットワークに IP マルチキャスト送信する．映像の復号に必要な暗号ヘッダ Hdr や，映像自身の画面サイズやタイトルといったメタ情報 I も同様に IP マルチキャスト送信する．これらの情報はマルチキャストに途中から参加するユーザのことも考慮する必要があるため，一定の配布間隔 t_i を置いて定期的に送信する．

このとき，ライブ映像配信の継続中にユーザが海賊行為を行ってユーザ鍵を流出させている可能性がある．ユーザ i の鍵 U_i の流出が検出された場合，そのユーザの鍵を無効化させなければならない．特定ユーザ鍵の無効化を実現するため，発信システムは対称鍵 K と暗号ヘッダ Hdr の定期的な更新を行い，常に有効ユーザ集合 A の変更が行えるようにする．更新には一定の計算量が必要となるため，リアルタイム配信に影響を及ぼさない範囲で可能な限り短い更新間隔 t_r での更新を行う．

一方，受信システムは BGW 復号化モジュール，SRTP 復号化モジュール，映像伸長及び電子指紋挿入モジュールからなる (図 4)．ユーザ i の BGW 復号化モジュールは定期的に受信する Hdr と，予め配布された U_i を元に対称鍵 K を復号する．SRTP 受信モジュールは K をもとに，受信した暗号化ストリーム C_M を復号し，映像ストリーム M を出力する．映像伸長及び電子指紋モジュールはメタ情報 I と M から映像を伸長し，同時にユーザ IDid_i を電子透かしとして挿入する．

実際の構築にあたり，映像圧縮モジュールと映像伸長モジュールには Windows Media Encoder 及び Windows Media Player を用い，映像圧縮符号化形式には Windows Media Video 9 形式を用いた．BGW 暗号化モジュール及び SRTP 暗号化モジュールは，Windows Meida Encoder に HTTP により映像ストリームを受け取り SRTP により IP マルチキャスト送信を行う送信プロキシとして実装した．BGW 復号化モジュール及び SRTP 復号化モジュールは，SRTP により暗号化ストリームを受けとり，HTTP により Windows Media Player に出力する受信プロキシとして動作する．SRTP による共通鍵暗号化の暗号方式には広く普及している AES を用いた．また，映像伸

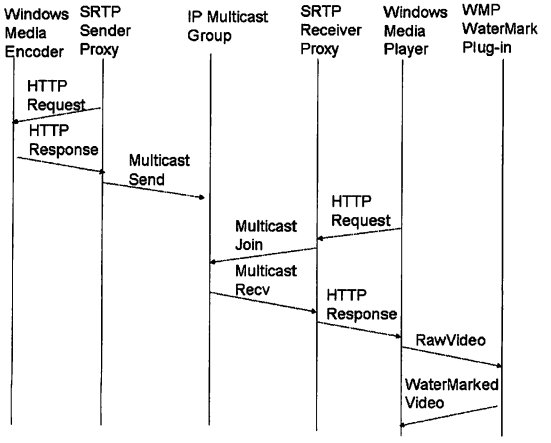

図 5 シーケンス図

図 6 映 像 例

表 1 フレームあたり SRTP 暗号化時間の平均 (sec)

	2M	1M	500k
送信	0.000031	0.000009	0.000002
受信	0.000056	0.000027	0.000020

長モジュールの電子指紋挿入は Windows Media Player のプラグインとして開発した．受信者による映像の配信要求開始から閲覧開始までのシーケンスを図 5 に示す．また，電子透かしの挿入には，PC 上でリアルタイムにソフトウェアによる透かしの挿入が可能な C4 Technology 社の M5 透かし方式を用いた．

4. 評価実験

提案システムを PC 上に構築し，計算資源と帯域資源の使用状況を評価する実験を行った．ライブ映像発信システム・受信システムの構築には，数十万円以内で市場で入手可能な PC を用いた．これらの PC は，Xeon 2.8GHz デュアル CPU と 2GB のメモリを搭載し，OS は Windows XP SP2 を用いた．対象とするユーザ数 n は数百人程度から数万人程度とし，比較的大規模な IP マルチキャストにも対応可能なシステムを想定している．

ライブ映像の撮影には Logitech Qcam 400 Pro を用いた (図 6)．撮影された大きさ QVGA(320x240)，毎秒 30 フレームの映像を，Windows Media Video 形式で圧縮した．このとき，圧縮映像の構造の最小単位である GOP(Group Of Picture) の時間長は 1(sec) とした．

表 2 対称鍵 K の暗号化，復号化にかかる時間 t_K^{gen}, t_K^{res} の平均値，標準偏差，最大値，最小値 (sec)

	平均値	標準偏差	最大値	最小値
t_K^{gen}	0.0419	0.0073	0.0781	0.0312
t_K^{dec}	0.0363	0.0079	0.0937	0.0156

表 3 フレームあたり電子指紋挿入時間の平均 (sec)

2M	1M	500k
0.0137	0.0133	0.0125

まず，SRTP 暗号化モジュール及び SRTP 復号化モジュールの計算資源の評価を行った．圧縮後のビットレートが 2Mbps, 1Mbps, 500Kbps における，SRTP 暗号化モジュール・復号化モジュールそれぞれの，1 フレームの映像データの処理に必要な CPU 使用時間を五分間計測し平均を算出したところ，表 1 のようになった．

SRTP は全ての映像ストリームを暗号化するので，秒間 30 フレームの映像のリアルタイム処理を行うためには映像 1 フレームあたりの計算時間が $\frac{1}{30}$sec を超えない必要がある．表 1 の結果，送受信時の 1 フレームあたりの処理時間は $\frac{1}{30}$sec の 1%にも満たないため，単体では十分にリアルタイム処理が可能であることが言える．

次に，BGW 暗号化モジュール・復号化モジュールの対称鍵の BGW 暗号化・BGW 復号化に必要な計算資源の評価を行った．BGW 暗号化・復号化で計算量が大きいのはペアリング計算 [6] と整数群のべき乗計算である．用いたペアリング計算は Pentium III のマシンで 10-100msec 程度で計算できることが報告されている [7]．べき乗計算は binary algorithm により指数 n に対して O(log$n$) で計算できることが知られているが，用いる底が 10^{76} 程度，指数が 10^{47} 程度とともに巨大で計算機の通常の整数演算の範囲に収まらないため，定性的な評価が難しい．従って，今回は実機による計算時間の評価を行なった．BGW 暗号化・復号化に要した CPU 使用時間 t_K^{gen},t_K^{res} を 100 回計算を繰り返して行い，平均値・標準偏差・最大値，最小値を算出したところ，結果は表 2 の通りであった．BGW 放送型暗号に基づく手法では，BGW 暗号化・BGW 復号化に要する時間はユーザ数 n とは関係なく一定である．

また，電子指紋の挿入モジュールが電子指紋の挿入に要した CPU 使用時間は表 3 の通りであった．

上記処理のうち，SRTP による暗号化・復号化及び電子指紋の挿入に比べ，放送暗号モジュールの対称鍵の BGW 暗号化・復号化には比較的大きな時間がかかるため，処理をフレームの暗号化処理のバックグラウンドで行う必要がある．両処理をあわせて CPU を使い切るとフレーム落ちなどの処理落ちを引き起すため，リアルタイム性を保持するためには対称鍵の更新間隔 t_r を大きめに設定する必要がある．

CPU 使用時間を基準に考えると，発信システムでは一つの

表 4　対称鍵 K の更新間隔 t_r に対し，送信システム，受信システムそれぞれが必要とする下限 (sec)

	2M	1M	500k
送信	0.042	0.042	0.042
受信	0.021	0.022	0.023

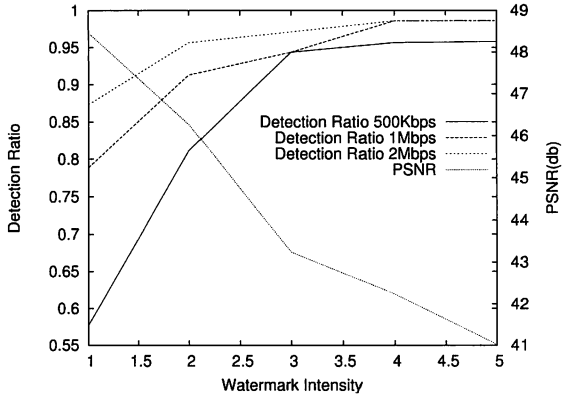

図 7　透かし強度変化に対する PSNR(db) 及びの指紋検出率の変化

鍵の生成に必要な CPU 時間を t_K^{gen}(sec)，1 フレームの暗号化に必要な時間が t_{enc} とすると，以下の式を満たす必要がある．

$$\frac{t_K^{gen}}{t_r} \times \frac{1}{30} + t_{enc} < \frac{1}{30} \tag{1}$$

また，受信システムでは鍵の BGW 復号化に必要な CPU 時間を t_K^{res}(sec)，1 フレームの復号化に必要な時間が t_{dec}，1 フレームの電子指紋挿入に必要な時間 t_f とすると，以下の式を満たす必要がある．

$$\frac{t_K^{res}}{t_r} \times \frac{1}{30} + t_{dec} + t_f < \frac{1}{30} \tag{2}$$

送受信の両方で処理落ちを防ぐためには，t_r は式 1, 2 を共に満す必要がある．2Mbps, 1Mbps, 500Kbps の映像ストリームにおいて表 1, 表 2, 表 3 の結果を式 1, 2 に代入した場合，送信，受信の双方が必要とする t_r の下限は表 4 のようになった．表 4 より，映像ストリームのビットレートにかかわらず送信側の方が長い更新間隔を必要としており，実験環境では t_r を 0.042(sec) 以上にすれば良いことがわかる．ただし，映像の圧縮や伸長は GOP の時間長を単位として行われるため，実際の更新間隔はこの条件を満しかつ GOP の時間長の整数倍とすることが妥当である．本実験では GOP の時間長が 1(sec) であるため，更新間隔 t_r を 1(sec) として配信実験を行った．

また，電子指紋挿入に伴う画質劣化，及び透かしの圧縮耐性の評価を行うため，電子指紋挿入後の PSNR と，伸長した透かし入り映像を 2Mbps, 1Mbps, 500Kbp それぞれに再圧縮を行った際の指紋検出率を図 7 に示す．利用した電子透かしモジュールでは透かし強度を五段階に設定可能であり，強度を強くするほど圧縮耐性は強まるが画質が劣化する．これらはトレードオフの関係にあるため，ここでは各透かし強度に対し (PSNR * 検出率) を評価関数としこれを最大化する強度を用いることとする．図 7 のデータにこの評価関数を適用した結果，2Mbps, 1Mbps, 500Kbps いずれの圧縮率においても透し強度 2 が最良と判断されるため，本実験ではこれを強度として用いた．ただし，受信側の計算機環境が非力である場合，電子指紋の挿入対象は全てのフレームではなく確率的に選択されたフレームにせざるを得ないため，そのような環境では画質劣化を許容し，強度の強い電子透かしを挿入する必要がある．

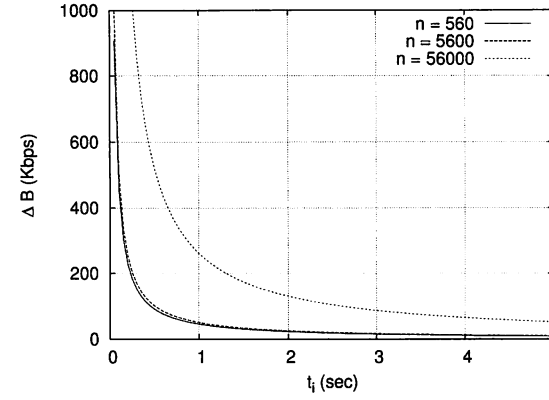

図 8　$n = 560, 5600, 56000$ のときのメタ情報及び暗号ヘッダの配布間隔 t_i に対する使用帯域の増加 ΔB(Kbps)

次に，帯域資源に対するオーバーヘッドの評価を行う．提案システムの適用が使用帯域に与える影響として IP マルチキャスト用のメタ情報 I のデータ量 D_I(bit)，及び暗号ヘッダ Hdr の配送による鍵のデータ量 D_{Hdr}^n(bit) がある．暗号ヘッダのデータ量 D_{Hdr}^n はユーザ数 n によって変化する．メタ情報及び暗号ヘッダの配布間隔を t_i(sec) とすると，使用帯域に与える影響 ΔB は

$$\Delta B = (D_I + D_{Hdr}^n)/t_i \tag{3}$$

提案システムにおいて，$n = 560, 5600, 56000$ の場合に D_I, D_{Hdr} と t_i の関係を求めたところ図 8 のようになった．例えば $n = 5600$ の場合，$1Mbps$ のストリームに対し使用帯域の増加を 10%以内に抑えたい場合，t_i を 0.51(sec) 以上に設定することで条件を満す使用帯域の設定が可能となる．本実験では，対称鍵 K の更新時間と同様，GOP の時間長の整数倍となるよう配布間隔 t_i を 1(sec) として配信実験を行った．

5. おわりに

我々は IP マルチキャスト上の SRTP を用いた鍵配送に放送型暗号を用いたライブ映像配信システムを，ソフトウェアによる実装で構築した．実験では数十万円以内で市場で入手可能な PC 上でソフトウェアを動作させ，数百人から数万人のユーザ数を想定し，帯域資源と計算資源のオーバーヘッドを計測した．結果，対称鍵の更新間隔とメタ情報及び暗号ヘッダの配布間隔を 1 秒に設定することで，処理落ちが無くかつ帯域増加を 10%以内に抑え，ライブ映像配信に十分に適用可能なことを示

した．今後は構築したシステムを実際にライブ映像配信を希望するユーザに利用してもらい，リアルタイム性や画質に関する主観的な評価を行う予定である．ライブ映像配信のアプリケーションとしては，遠隔教育のための講義映像配信などを想定している．

文　献

[1] D. Boneh, C. Gentry, B. Waters, Broadcast Encryption, Proc. of Crypto '93, LNCS 773, pp. 480–491, 1993

[2] A. Fiat, M. Naor, Collusion Resistant Broadcast Encryption with Short Ciphertexts and Private Keys, Proc. of Crypto '05, LNCS 3621, pp. 258–275, 2005

[3] B. B. Amberker, P. Koulgi, M.B.Nirmala, Some Implementation Issues in the Fiat-Naor Broadcast Encryption Schemes, Proc. of ADCOM 2006, pp.634–635, 2006

[4] I. Echizen, K. Tanimoto, T. Yamada, M. Daninaka, S. Tezuka, H. Yoshiura, PC-based real-time watermark embedding system with standard video interface, Proc. of SMC2006, pp. 267–271, 2006

[5] M. Baugher, et al., The Secure Real-Time Transport Protocol (SRTP), http://www.ietf.org/rfc/rfc3711.txt, 2004

[6] Ben Lynn, On The Implementation On Pairing-Based Cryptosystems, Stanford University Ph. D. Thesis, p.111, 2007

[7] Ben Lynn, The Pairing-Based Cryptography Library Benchmarks, http://crypto.stanford.edu/pbc/times.html