멀티캐스트 환경에서 효율적인 그룹키 관리를 위한 트리구조 및 알고리즘 개발

한 근 희†

요 약

멀티캐스트 환경에서 그룹키 관리는 그룹 통신에서 기밀성과 같은 정보보호서비스를 제공하기 위하여 그룹 내 모든 구성원들이 한 개의 동일한 비밀키를 공유한 후 이를 기반으로 그룹 내 메시지의 암복호화를 수행함으로서 그룹 통신 내용을 보호하는 기술이다. 지금까지의 그룹키 관리 연구에서는 완전 정규 트리를 기반으로 사용자들이 그룹에 가입 및 탈퇴하는 비율을 고려하지 않은 상태에서 그룹키 관리 메커니즘이 개발되어 왔지만 본 연구에서는 사용자들이 그룹에 가입 및 탈퇴하는 비율을 고려한 상태에서 관리 메커니즘의 효율성을 높이는 메커니즘을 개발하였으며 또한 완전 정규 트리 보다 더욱 유연한 구조를 갖는 가변트리라는 새로운 트리구조를 정의 및 분석하여 제안된 가변트리 모델이 완전 정규트리 모델보다 그룹키 관리에서 더욱 효율적인 키트리 모델임을 제시하였다. 특히 그룹 통신에서 탈퇴비율이 50%를 넘는 경우 트리의 차수들이 2 또는 3인 경우에 최적화가 이루어짐을 증명하였다.

Development of Tree Structures and Algorithms for the Efficient Group Key Management in Multicast Environment

Keunhee Han†

ABSTRACT

In multicast environment, the main objective of group key management is to provide security services to group communications by sharing a single group key among all the members of the group and subsequently encrypting and decrypting all the communication messages exchanged among the members of the group. Up to now, there has been no effort to develop group key management mechanism that considers the rate of users' join/leave operations. Hence, in this research, we propose group key management mechanisms that consider the rate of user's join/leave operations. We also define a new tree structure called variable tree which is much more flexible than full regular trees and show that variable trees are more efficient than full regular trees for group key management. Especially, we propose an algorithm that minimizes the necessary number of rekey messages according to the rate of join and leave operations. We also shows that if the rate of leave operation is greater than 50%, then the tree structure with degrees 2 or 3 are the optimal structures.

키워드 : 그룹키 관리, 멀티캐스트, 그룹통신, 정보보호, 알고리즘

1. 서 론

멀티캐스트 통신에서 그룹 내 사용자들 사이에 전송되는 메시지를 보호하는 것은 단대단 통신의 경우와 마찬가지로 메시지 암호화를 통하여 이룰 수 있다. 그러나 단대단 통신과는 달리 멀티캐스트 통신에서 메시지를 효율적으로 암복호화하기 위해서는 그룹 내의 모든 사용자들이 한 개의 비밀키를 동일하게 공유하여야 한다는 문제점이 있다. 대개의 경우 멀티캐스트 그룹은 동적인 그룹으로서 기존의 가입자가 그룹을 탈퇴할 수도 있으며 또한 새로운 사용자가 그룹에 가입할 수도 있다. 기존의 가입자가 그룹을 탈퇴하는 것을 탈퇴 동작(Leave operation) 이라 하며 새로운 사용자가 그룹에 가입하는 것을 가입 동작 (Join operation) 이라 한다. 강한 정보 보호 서비스를 제공하기 위해서는 이러한 사용자들의 가입 및 탈퇴 동작이 발생할 때마다 PFS(Perfect Forward and Backward Secrecy) 라는 정보보호 요구사항을 만족시켜야만 한다.

PFS는 그룹 내 사용자 중 그룹을 탈퇴하는 가입자가 그룹을 탈퇴한 이후의 통신내용을 인지할 수 없어야 하는 것과 또한 그룹에 새로이 가입하는 사용자가 가입 이전의 통신내용을 인지할 수 없어야 한다는 정보 보호 요구 조건이

※ 본 연구는 정보통신부 정보통신연구진흥원에서 지원하고 있는 대학 기초지원연구사업(과제번호 2001-123-2)에 의하여 지원되었음.
† 종신회원 : 공주대학교 응용수학과 교수
논문접수 : 2002년 7월 27일, 심사완료 : 2002년 10월 2일

다. PFS 요구조건을 충족시킬 수 있는 방법으로서는 멀티캐스트 그룹에 가입 및 탈퇴 동작이 발생할 때마다 기존에 모든 가입자들 사이에 공유되고 있는 그룹키를 새로운 그룹키로 대체하는 것이 유일한 방법이다. 즉, 가입 및 탈퇴 동작이 발생할 때마다 기존의 그룹키를 새로운 그룹키로 대체함으로서 탈퇴하는 사용자는 탈퇴한 이후의 그룹 메시지를 복호화할 수 없고 새로이 가입하는 사용자는 가입 이전에 비록 특정한 방법을 통하여 그룹 메시지를 수집 및 저장하였다 하여도 이들을 복호화할 수 없도록 하는 것이다.

그룹키 관리 연구는 크게 중앙 집중적 방식 및 분산 관리 방식으로 분류될 수 있으며 지금까지 IRTF 내의 SMuG (Secure Multicast) 작업반에서 가장 활발히 진행되어 왔으나 최근에는 IETF 내에 Msec (Multicast Security) 이라는 그룹키 관리 및 멀티캐스트의 정보 보호 연구를 담당하는 작업반이 결성되어 SMuG 및 RMRG (Reliable Multicast) 작업반 등과 함께 본 분야의 연구를 주도하고 있다.

분산 관리 방식의 연구는 Hardjono[4]등에 의한 Intra-domain Group key Management Protocol이 제안되었으나 분산관리 방식은 그룹키 관리를 위한 데이터 구조 및 알고리즘의 개발보다는 관련 프로토콜 개발에 연구가 집중되고 있다. 이는 분산 방식이 복수개의 키 관리 서버들이 그룹키를 관리하므로 효율적인 데이터 구조 및 알고리즘보다는 대량의 사용자들과 지역적으로 효율적으로 통신할 수 있는 기술의 개발을 추구하기 때문이다. 분산 관리 방식의 다른 연구로서는 Suvo Mittra에의한 Iolus[5]를 들 수 있다. 본 방식은 확장성을 염두에 두고 GSA(Group Security Agents)라는 중간 단계의 그룹키 서버들을 사용한 메커니즘이다.

분산 관리 방식에 비하여 중앙 집중 관리 방식의 가장 큰 특징은 중앙 집중 방식에서는 한 개의 그룹키 서버가 모든 그룹키들을 관리한다는 것이다. 따라서, 중앙 집중 방식은 분산 관리 방식보다 정보 보호 관점에서 훨씬 강한 안전성을 제공할 수 있다. 중앙 집중 관리 방식의 가장 대표적인 데이터 구조 및 알고리즘은 LKH(Logical Key Hierarchy)[2]를 들 수 있으며 트리(Tree)를 기반으로 한 데이터 구조를 제시하고 있다. 다른 중앙 집중적 그룹키 관리 방식으로서 OFT(One-way Function Tree)[6]라는 메커니즘이 Balenson에 의하여 제안되었으며 Moyer[3]는 트리 구조의 균형을 유지하는 방안을 연구하였다. 트리 구조를 이용한 또 다른 키관리 방식으로서 Wong[1]의 키트리 모델이 있다. 본 방식은 LKH와 유사하지만 LKH와는 달리 사용자 중심 모드(User-oriented), 키 중심 모드(Key-oriented) 및 그룹중심 모드(Group-oriented) 등 3가지 모드로 그룹키 정보를 구성하여 사용자들에게 전송할 수 있도록 개발되었다.

새로운 사용자 u가 그룹 통신에 가입 동작을 요청하면 서버와 사용자는 IKE[7]등 적절한 인증 프로토콜을 이용하여 상호간의 신원 인증 과정을 거친 후 그룹키 서버 S와 u 사이에만 공유되는 개인키를 설정하게 된다. 이와 같이 그룹키 서버와 개개 가입자 사이에 개인키를 설정하는 과정을 초기화 과정이라 하며 만일 그룹 통신에 n명의 사용자가 가입한다면 모두 n번의 초기화 과정이 필요하다. 서버 S는 u의 개인키로 그룹키를 암호화하여 u에게 전송함으로서 가입동작은 완료되며 새로운 사용자 u는 전송 받은 그룹키를 이용하여 멀티캐스트 통신에 참여할 수 있게 된다. 기존의 가입자가 그룹을 탈퇴하는 경우 서버는 탈퇴하는 사용자와 공유되었던 개인키를 서버에서 관리하는 데이터 구조에서 삭제함으로서 탈퇴동작이 완료될 수 있을 것이다.

사용자들의 가입 및 탈퇴 동작이 발생할때 PFS 요구조건을 만족시키기 위하여 그룹키 서버가 기존에 모든 사용자들 사이에 공유되던 그룹키를 새로운 그룹키로 대체하는 과정을 "키갱신"이라 한다. 이 과정에서 서버는 사용자들에게 새로이 생성된 그룹키를 적절한 비밀키 k로 암호화하여 저장한 키갱신 메시지(rekey message)를 전송하여야 한다. 멀티캐스트 그룹 내에 n명의 사용자가 있는 경우 만일 서버와 가입자들 사이에 초기화 과정에 개별적으로 설정된 개인키를 k로 사용하여 단대단 통신을 이용하여 키갱신 메시지를 전송한다면 이러한 과정에는 명백히 $O(n)$개의 키갱신 메시지가 전송된다. 만일 n이 큰 수이며 빈번한 가입/탈퇴 동작이 발생한다면 단대단 통신을 이용한 키갱신 메커니즘은 과도한 통신량으로 인하여 멀티캐스트 통신의 근본적인 장점이 상쇄되게 된다.

키갱신을 효율적으로 수행하기 위하여 Wong[1] 및 Wallner[2] 등에 의하여 트리(Tree) 구조가 제시되었지만 이들에 의하여 제시된 트리는 완전 정규 트리 이기 때문에 그룹 통신에 참여하는 사용자 수에 따라 알맞은 크기의 트리 구조를 제시하는데 어려움이 있다. 또한 동적인 멀티캐스트 그룹에서 사용자의 가입/탈퇴 비율에 따라 더욱 효율적인 트리 모델을 제시할 수 있는 연구가 현재까지 알려지지 않았다.

따라서, 본 연구에서는 구조적인 면에서 완전 정규 트리보다 훨씬 유연한 가변 트리라는 새로운 트리 구조를 정의 및 분석하여 이들을 기반으로 한 그룹키 관리의 최적화 방안을 연구하였으며 특히 가입 및 탈퇴 동작의 발생 비율을 감안한 상태에서 최적의 트리 구조를 분석하는 방안이 연구되었다. 또한 사용자의 가입 및 탈퇴 동작들을 개별적으로 취급하였을 경우 이들 각각에 대한 최적의 트리 모델들을 제시 및 증명하였다. 특히, 그룹 통신에서 사용자의 가입/탈퇴 비율을 충분히 예측할 수 없는 경우를 고려하여 가장 일반적인 경우, 즉 가입/탈퇴 비율이 각각 50%인 경우에 사용될 수 있는 최적의 트리 구조를 제시하였다.

본 논문의 나머지 부분은 다음과 같이 구성되어 있다. 2

장에서는 Wong[1]에 의하여 기존에 제시된 키트리 모델 및 키갱신 알고리즘을 분석하며 3장에서는 먼저 키트리 모델에 대한 분석을 통하여 사용자의 가입 및 탈퇴 동작에서 최적화된 키트리 구조들은 서로 상반된 구조를 갖는다는 것을 증명하며 이를 기반으로 사용자의 가입 및 탈퇴 비율을 고려한 상태에서 완전 정규 트리를 이용한 키트리의 최적화 방안을 연구한다. 4장에서는 완전 정규 트리에 내재된 구조상의 경직성을 피하기 위하여 새로이 가변트리라는 트리 구조를 정의하여 이를 기반으로 사용자의 가입 및 탈퇴 비율을 고려한 상태에서 가변트리를 이용한 그룹키 관리의 최적화 방안을 연구한다. 끝으로 5장에서는 결론을 맺는다.

본 논문에서는 다수의 그래프 용어가 사용되며 특별히 정의되지 않은 그래프 용어는 Harray[8]의 표준적인 그래프 용어들을 따르도록 한다.

2. 키트리 모델을 이용한 키갱신 메커니즘

본 절에서는 Wong[1]에 의하여 제시된 키트리 모델 및 관련 키갱신 알고리즘들을 설명하여 이를 기반으로 3장 및 4장에서는 각각 완전 정규 트리 및 가변트리를 기반으로 [1]에서 제시된 키트리 모델을 최적화할 수 있는 방안을 연구하도록 한다.

특정한 키갱신 메커니즘을 G라 하자. G를 이용한 키갱신 메커니즘에서 G의 효율성을 측정하기 위하여 다음과 같은 파라미터들이 고려되어야 한다.

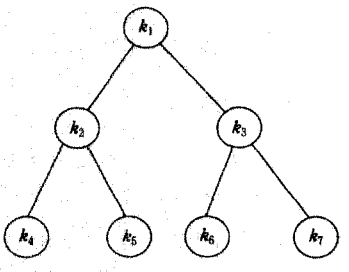
- JMsg(G) (LMsg(G)) : 한 개의 가입(탈퇴) 동작에 따른 키갱신 과정에서 서버가 사용자들에게 전송해야 되는 총 키갱신 메시지 갯수. 멀티캐스트 메시지는 한 개의 메시지로 계산된다.
- JEnc(G) (LEnc(G)) : 한 개의 가입(탈퇴) 동작에 따른 키갱신 과정에서 서버가 암호화해야 할 총 메시지 개수.
- SubKey(G) : G에 포함된 서브 그룹키의 총 개수.

멀티캐스트 통신의 근본 목적이 통신량의 최소화에 있기 때문에 상기한 효율성 파라메타들 중 가장 중요한 것은 JMsg(G) 및 LMsg(G)이라 할 수 있다. 따라서, 앞으로 최적의 키트리 모델이란 키갱신 과정에서 총 키갱신 메시지 개수를 최소화하는 모델을 의미한다.

키트리는 Wong[1]에 의하여 정의된 데이터 구조로서 일반 그래프의 구조를 갖지만 본 논문에서는 다음과 같이 키트리를 완전 정규 트리 구조로 한정하여 수정 정의한다.

[정의 2.1] "키트리"는 완전 정규 트리(Full Regular Tree)로서 말단 정점들을 제외한 모든 정점들은 동일한 개수의 자식 정점을 갖는다. 키트리의 루트는 그룹키 서버 역할을 하며 말단 정점들은 가입자의 역할을 한다. 루트 및 말단 정점들을 제외한 모든 내부 정점들은 서브 그룹키(Subgroup key)의 역할을 한다. 키트리 T의 차수 d는 T의 내부정점이 갖는 자식정점의 개수이며 T의 높이 h는 루트로부터 말단정점에 이르는 경로 상의 정점의 개수이다. T에서 루트의 레벨을 1로 정의하며 나머지 정점들은 루트의 레벨로부터 1씩 증가한다.

(그림 1)의 키트리 T는 $d=2$ 및 $h=3$이며 모두 7개의 정점으로 구성되어 있으며 루트에 포함된 k_1은 그룹키이며 k_2 및 k_3는 서브 그룹키로서 키갱신 과정에서 효율적으로 기존의 그룹키를 갱신하는데 사용된다. $k_4 \sim k_7$은 서버와 가입자들 사이에만 개별적으로 공유되는 개인키이다. 키트리에서 u_i는 k_i의 개인키를 갖는 가입자(말단정점)를 나타낸다.



(그림 1) 키트리 T

Wong[1]은 키갱신 과정을 위하여 사용자 중심 모드, 키 중심 모드 및 그룹 모드 등 모두 3가지의 서로 다른 모드를 제시하였지만 이들 중 키 중심 모드가 효율성에서 가장 뛰어나므로 본 연구에서는 키 중심 모드를 키갱신의 기본 알고리즘으로 채택하며 다른 2가지 모드에 대해서는 [1]을 참조하도록 한다. 그룹키 갱신 과정에서 서버의 메시지 전송 프로토콜을 다음과 같이 정의한다.

① $S \rightarrow u_i : m$
서버 S가 사용자 u_i에게 메시지 m을 단대단 통신으로 전송하는 것.

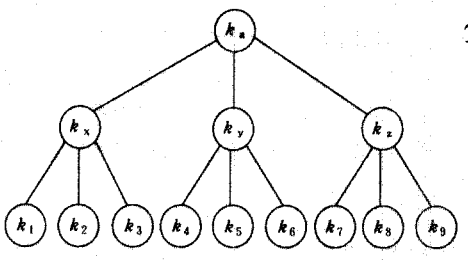
② $S \rightarrow \{u_1, u_2, \cdots, u_n\} : m$
서버 S가 사용자 집합 $\{u_1, u_2, \cdots, u_n\}$에게 메시지 m을 멀티캐스트를 이용하여 전송하는 것.

③ $\{k_i\} k_j$
키 k_i를 키 k_j로 암호화한 메시지.

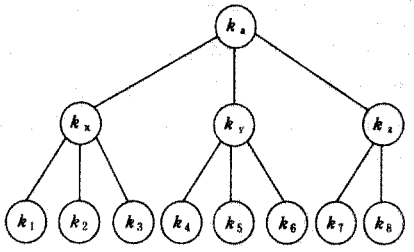
④ k_i'
키 k_i를 대체하기 위하여 새로이 생성된 키

(그림 2)의 키트리 T는 (그림 3)의 키트리 T'로부터 한 명의 새로운 가입자가 새로이 가입동작을 통하여 가입한 후의 키트리를 나타내며, 이때 새로운 가입자는 서버로부터 k_9의 개인키를 부여받게 된다. 키트리 T'는 T로부터 k_9의

개인키를 가진 기존의 가입자가 탈퇴한 후의 키트리를 나타낸다.



(그림 2) 키트리 T



(그림 3) 키트리 T'

(그림 3)의 키트리 T'에 u_9이 새로이 가입한다고 하자. 서버 S는 u_9에게 k_9을 부여하여 T'를 T로 수정한 후 PFS를 위하여 u_9의 부모정점부터 루트에 이르는 경로상에 위치한 모든 서브 그룹키 및 그룹키들을 새로운 키들로 생성한 후 다음과 같은 키갱신 메시지들을 가입자들에게 전송함으로서 가입동작이 완료된다.

$S \rightarrow \{u_1, \cdots, u_6\}$: $\{k'_a\}k_a$
$S \rightarrow \{u_7, u_8\}$: $\{k'_a\}k_a$, $\{k'_z\}k_z$
$S \rightarrow u_9$: $\{k'_a, k'_z\}k_9$

본 과정에서 JEnc(T) = $2(h-1)$이 요구된다. 이는 상기한 예제에서도 나타나듯이 동일한 키로 암호화된 키들이 중복되게 사용될 수 있기 때문이다. 예를 들어, $\{k'_a\}k_a$은 $\{u_1, \cdots, u_6\}$, 및 $\{u_7, u_8\}$에게 반복적으로 전달되고 있기 때문에 $\{k'_a\}k_a$는 한번만 생성되면 복수의 가입자들에게 재 전송이 될 수 있다. JMsg(T) = h이며 이는 기존의 사용자들을 위하여 새로운 그룹키 및 서브 그룹키들이 한번씩 만 암호화되어 전송되며 (따라서, $h-1$개의 키갱신 메시지), 또한 u_9에게는 이러한 $h-1$개의 키들을 하나의 메시지로 묶은 후 k_9으로 암호화하여 전송하기 때문이다.

(그림 2)의 키트리 T로부터 u_9이 그룹을 탈퇴하게 되면 서버 S는 T로부터 u_9을 삭제하여 키트리를 T'로 수정하며 가입 동작과 마찬가지로 u_9의 부모 정점으로부터 루트에 이르는 모든 서브 그룹키 및 그룹키를 새로운 키들로 대체 생성한 후 다음과 같은 키갱신 메시지들을 그룹에 남아 있는 사용자들에게 전송함으로서 탈퇴 동작을 완료한다.

$S \rightarrow \{u_1, u_2, u_3\}$: $\{k'_a\}k_x$
$S \rightarrow \{u_4, u_5, u_6\}$: $\{k'_a\}k_y$
$S \rightarrow u_7$: $\{k'_a\}k'_z$, $\{k'_z\}k_7$
$S \rightarrow u_8$: $\{k'_a\}k'_z$, $\{k'_z\}k_8$

본 과정에서 LMsg(T) = $(d-1)(h-1)$이 요구되며 이는 탈퇴하는 사용자의 부모 정점으로부터 루트에 이르는 $(h-1)$개의 키들이 자신의 자식 노드에 포함된 서로 다른 서브 그룹키로 암호화된 후 전송되기 때문이다. 또한 본 과정에서 키들이 암호화되는 과정은 가입 동작의 경우와는 다소 상이하다. 예를 들어, 상기한 프로토콜에서 사용자 u_7는 암호문 $\{k'_z\}k_7$을 자신의 개인키 k_7을 복호화하여 k'_z를 얻은 후에 $\{k'_a\}k'_z$를 복호화할 수 있으며 이것은 u_8도 마찬가지이다. [1]에서 LEnc(T) = $d(h-1)$이라고 계산되어 있지만 실제로는 다음에 증명된 것과 같이 $d(h-1)-1$이다.

[정리 2.1] n명의 사용자를 포함한 키트리 T의 높이가 h이고 차수가 d면 키 중심 모드의 탈퇴 동작에서 LEnc (T) = $d(h-1)-1$이다.

(증명) 탈퇴 동작의 경우 T에서 탈퇴하는 사용자 정점의 부모 정점부터 루트까지의 경로 상에 위치한 $(h-1)$개의 키들이 새로이 생성된다. 이들 중 루트부터 레벨 $(h-2)$ 사이의 $(h-2)$개의 키들은 자신들의 서브 그룹키들로 각각 암호화된다. 따라서 $d(h-2)$번의 암호화가 이루어진다. 레벨 $(h-1)$ 즉 탈퇴하는 사용자의 부모 정점에 위치한 새로이 생성된 키는 자신의 자식 노드에 있는 $(d-1)$개의 사용자들의 개인키로 암호화가 된다. 따라서, 다음이 성립한다.

$$\text{LEnc}(T) = d(h-2) + (d-1) = d(h-1) - 1. \quad \blacksquare$$

3. 완전 정규 트리를 이용한 그룹키 최적화

본 절에서는 [1]에서 제시된 키트리 모델에 대한 자세한 분석을 통하여 이를 최적화할 수 있는 방안을 연구하도록 한다.

2 절에서 키트리 T의 효율성은 T의 높이 h 및 차수 d에 의하여 결정된다는 것을 알 수 있다. 다음 정리는 완전 정규 트리의 높이 h 값을 결정한다.

[정리 3.1] 최소한 n개의 말단 정점을 수용할 수 있는 차수 d를 갖는 완전 정규 트리 T의 높이 h는 다음 식을 만족하여야 한다.

$$h \geq \lceil \log_d n \rceil + 1 \qquad (3.1)$$

(증명) 만일 $n = d^k (k = 1, 2, \cdots)$이면 명백히 $h = \log_d n + 1$이다. $n (\neq d^k)$을 $d^k < n < d^{k+1}$의 범위에 있는 정수라 하고 $h < \lceil \log_d n \rceil + 1$이라 하자. 그렇다면 정수 h의 최대값은 이 된다. 또한 $d^k < n < d^{k+1}$이므로 $\lceil \log_d n \rceil = k+1$이다. 완전 정규 트리의 높이가 $k+1$이라면 말단 정점의 개수는 d^{k+1-1}

$=d^k(<n)$이므로 $n(d^k < n < d^{k+1})$개의 말단 정점을 포함할 수 있는 완전 정규 트리의 높이 h는 식 (3.1)를 만족하여야 한다. ■■■

다음의 [부속정리 3.1]~[부속정리 3.5] 및 [정리 3.2]는 키트리 모델에서 탈퇴 동작의 경우 $d = 2$ 일 때 즉, 키트리가 완전 이진 트리인 경우에 가장 적은 수의 키갱신 메시지를 요구한다는 것을 보여 준다.

[부속정리 3.1] x 및 y를 2개의 실수라 하자. 만일 $x < y$ 라면 $\lceil x \rceil \le \lceil y \rceil$ 이다.

(증명) $x = a + r_1$ 및 $y = b + r_2$ (a, b는 정수, $0 r_1, r_2 < 1$)라 하자.

(경우 1) $r_1 = r_2 = 0$인 경우, $x = a$, $y = b$이며 $a < b$이므로 $\lceil x \rceil = a < b = \lceil y \rceil$ 이다. (경우 2) $r_1 = 0, r_2 \neq 0$인 경우, $a < \lceil b + r_2 \rceil = b + \lceil r_2 \rceil = b + 1$로부터 $a \le b$이므로 $\lceil x \rceil = a \le b < b + \lceil r_2 \rceil = \lceil b + r_2 \rceil = \lceil y \rceil$ 이다. (경우 3) $r_1 \neq 0, r_2 = 0$인 경우, $a + r_1 < b$로부터 $a + 1 \le b$이므로 $\lceil x \rceil = a \lceil r_1 \rceil = a + 1 \le b = \lceil y \rceil$ 이다. (경우 4) $r_1 \neq 0, r_2 \neq 0$인 경우, $a + r_1 < b + r_2$로부터 $a \le b$이므로 $\lceil x \rceil = \lceil a + r_1 \rceil = a + 1 \le b + 1 = b + \lceil r_2 \rceil = \lceil y \rceil$ 이다. ■■■

[부속정리 3.2] 정수 $n(>1)$및 $d(3 \le d \le n)$에 대하여 $\log_2 n < (d-1)(\log_d n)$이다.

(증명) $\log_2 n < (d-1)\log_d n$

$\rightarrow 1 < (d-1)/\log_2 d$ ($\because \log_2 n \neq 0$)

$\rightarrow 1 < (d-1)/\log_2 d < (d-1)$ ($\because \log_2 d > 1$)

$\rightarrow 1 < (d-1)$ ($d \ge 3$ 이므로 항상 참) ■■■

[부속정리 3.3] 정수 x 및 실수 $y(\ge 0)$에 대하여 $\lceil xy \rceil \le x \lceil y \rceil$ 이다.

(증명) $y = k + r$ (k는 정수, $0 \le r < 1$)이라 하자. 만일 $r = 0$이면 명확히 $\lceil xy \rceil = x \lceil y \rceil$ 이다. $r \neq 0$이면 $x \cdot r < x$ 이므로 $\lceil xr \rceil \le x = x \lceil r \rceil$ 이다. 따라서, $\lceil xy \rceil = \lceil x(k+r) \rceil = \lceil xk + xr \rceil = xk + \lceil xr \rceil \le xk + x \lceil r \rceil = x(k + r \rceil) = xy$. ■■■

[부속정리 3.4] 정수 $d = 3, \cdots, n$에 대하여 $\lceil \log_2 n \rceil \le (d-1) \lceil \log_d n \rceil$ 이다.

(증명) [부속정리 3.1], [부속정리 3.2] 및 [부속정리 3.3]에 의하여

$\log_2 n < (d-1)\log_d n$

$\rightarrow \lceil \log_2 n \rceil \le \lceil (d-1)\log_d n \rceil$

$\rightarrow \lceil \log_2 n \rceil \le (d-1) \lceil \log_d n \rceil$ ■■■

[부속정리 3.4]로부터 우리는 다음 정리를 구할 수 있다.

[정리 3.2] n명의 사용자를 포함하는 키트리 모델 T는 차수가 2일 때 탈퇴동작에서 최소한의 키갱신 메시지가 발생한다.

모든 키트리 모델의 가입동작에서 PFS를 위하여 기존에 사용되는 그룹키를 새로운 그룹키로 대체하여야 하며 특히 기존의 가입자와 새로이 가입하는 사용자에게 전송되는 키갱신 메시지는 PFS를 위하여 반드시 서로 다른 키로 암호화 되여야 하므로 모든 키트리 모델의 가입동작에는 최소한 2개의 키갱신 메시지가 필요하다. 따라서, 우리는 다음 정리를 얻는다.

[정리 3.3] n명의 가입자를 포함하는 키트리 모델 T는 차수 $d = n$ 일 때 가입동작에서 최소한의 키갱신 메시지가 발생한다.

(증명) $u_1 \sim u_n$ 등 모두 n명의 사용자를 포함하는 키트리 T의 차수를 n이라 하자. 그렇다면 T의 높이 h는 2이며 따라서 T에서의 가입동작에는 $h = 2$개의 키갱신 메시지가 필요하다. ■■■

[정리 3.2] 및 [정리 3.3]은 키트리 모델 T에서 가입 및 탈퇴 동작은 서로 상반된 구조에서 키갱신 메시지의 효율성이 최적화되고 있음을 보여 준다. 즉 가입 동작의 경우 T의 높이가 작을 수로 적은 수의 키갱신 메시지가 발생하지만 탈퇴 동작의 경우 T의 높이가 클 수록 적은 수의 키갱신 메시지가 발생한다는 것을 보여 준다. 이것이 키트리 모델의 최적화를 이루는데 가장 어려운 부분이다. 이러한 분석은 자연스럽게 키트리를 기반으로 한 그룹키 관리 구조에서 가입 및 탈퇴 동작이 발생하는 비율을 고려한 상태에서 키트리의 최적화를 이룰 수 있는 방안을 연구하여야 함을 제시하고 있다.

키트리 모델 T에서 가입 및 탈퇴동작이 50 : 50의 비율로 발생한다고 가정한 상태에서 A를 가입 및 탈퇴동작에 의하여 발생하는 평균 키갱신 메시지 개수라 하면 A는 다음과 같다.

$$A = (h + (h-1)(d-1))/2 = d(h-1)/2$$
$$= (1/2)d \lceil \log_d n \rceil \ (\because h = \lceil \log_d n \rceil + 1)$$

우리는 $d = 2$ 또는 3일 때 A가 최소값을 갖는다는 것을 [부속정리 3.5]~[부속정리 3.7] 및 [정리 3.4]를 통하여 알 수 있다.

[부속정리 3.5] 양의 정수 $n(\ge 1)$에 대하여 다음 관계가 성립한다.

(1) 만일 $n = d^k$(k는 $n \ge 1$을 만족하는 정수) 라면 $d = 2, 3, \cdots$에 대하여 $\lceil \log_d (n+1) \rceil = 1 + \lceil \log_d n \rceil$ 이다.

(2) 만일 $n \neq d^k$(k는 $n \ge 1$을 만족하는 정수) 라면 $d = 2, 3, \cdots$에 대하여 $\lceil \log_d (n+1) \rceil = \lceil \log_d n \rceil$ 이다.

(증명) (1) $n = d^k$이면 $\lceil \log_d(n+1) \rceil = \lceil k + r \rceil$ $(0 < r < 1)$이다. 따라서, $\lceil \log_d(n+1) \rceil = \lceil k + r \rceil = 1 + k = 1 + \lceil \log_d d^k \rceil = 1 + \lceil \log_d n \rceil$. (2) n이 $d^{k-1} < n < d^k$이라 하자. 그렇다면, $\lceil \log_d n \rceil = \lceil (k-1) + r \rceil$ $(0 < r < 1)$ 이므로 $\lceil \log_d n \rceil = k$가 된다. 또한 만일 $n = d^k - 1$이라면 $\lceil \log_d(n+1) \rceil = \lceil \log_d d^k \rceil = k$이며 $n \neq d^k - 1$인 경우에도 $\lceil \log_d(n+1) \rceil = k$가 된다. 따라서, 모든 경우에서 $\lceil \log_d(n+1) \rceil = \lceil \log_d n \rceil$ 이다. ■■■

[부속정리 3.6] $n(\geq 6)$ 및 $4 \leq d \leq n$을 만족하는 모든 정수 n 및 d에서 대하여 $2\lceil \log_2 n \rceil \leq d \lceil \log_d n \rceil$ 이다.

(증명) 증명은 n에 대한 귀납법을 이용한다. $n = 6$이면 $2\lceil \log_2 6 \rceil = 6$이고 $4 \leq d \leq n$ $(= 6)$ 범위의 d값에 대하여 $4\lceil \log_4 6 \rceil = 8$, $5\lceil \log_5 6 \rceil = 10$ 및 $6\lceil \log_6 6 \rceil = 6$이므로 모든 경우에서 $2\lceil \log_2 n \rceil \leq d \lceil \log_d n \rceil$ 이다. 이제 n (> 6) 및 $4 \leq d \leq n$에 대하여 $2\lceil \log_2 n \rceil \leq d \lceil \log_d n \rceil$ 이 성립한다고 가정하고 $2\lceil \log_2(n+1) \rceil \leq d \lceil \log_d(n+1) \rceil$ 임을 증명하도록 한다. 우리는 n의 크기에 따라 다음과 같은 2가지 경우를 증명한다.

(1) $n = 2^k$ $(k \geq 3)$인 경우 :

$2\lceil \log_2(n+1) \rceil = 2(\lceil \log_2 n \rceil + 1)$ (∵ [부속정리 3.5](1))
$\leq d \lceil \log_d n \rceil + 2$ (∵ 귀납가정)
$= d(\lceil \log_d(n+1) \rceil - 1) + 2$ (∵ [부속정리 3.5](1))
$\leq d \lceil \log_d(n+1) \rceil$ (∵ $(d-2) \geq 2$)

(2) $n \neq 2^k$ $(k \geq 3)$인 경우 :

$2\lceil \log_2(n+1) \rceil = 2\lceil \log_2 n \rceil$ (∵ [부속정리 3.5](2))
$\leq d \lceil \log_d n \rceil$ (∵ 귀납가정)
$= d \lceil \log_d(n+1) \rceil$ (∵ [부속정리 3.5](2)) ■■■

[부속정리 3.7] 주어진 n값에 대하여 $A = (1/2)d \lceil \log_d n \rceil$은 $d = 2$ 또는 $d = 3$ 일 때 최소값을 갖는다.

(증명) [부속정리 3.6]에서 $d \geq 4$인 경우에 $2\lceil \log_2 n \rceil \leq d \lceil \log_d n \rceil$ 임을 알 수 있다. 따라서, 만일 $2\lceil \log_2 n \rceil \leq 3\lceil \log_3 n \rceil$ 이라면 $2\lceil \log_2 n \rceil$이 최소값이 되며 만일 $3\lceil \log_3 n \rceil < 2\lceil \log_2 n \rceil$ 라면 $3\lceil \log_3 n \rceil$이 최소값이 된다. 따라서 $A = (1/2)d \lceil \log_d n \rceil$은 $d = 2$ 또는 $d = 3$ 일 때 최소값을 갖는다. ■■■

[부속정리 3.7]로부터 A 값이 최소가 되는 정확한 d 값을 한번에 결정할 수는 없다. 그러나, $d = 2$ 또는 3 인 경우를 계산한 후 이들 값들을 단순 비교함으로서 A가 최소가 되는 정확한 d 값을 알 수 있으며 우리는 다음과 같은 정리를 얻는다.

[정리 3.4] 완전 정규 트리 모델 T에서 가입 및 탈퇴 동작이 50 : 50의 비율로 발생할 때 키갱신 메시지를 최소화할 수 있는 구조는 T의 차수 d가 2 또는 3일 때이다. ■■■

더욱 일반적인 경우를 고려하기 위하여 p를 키트리 모델에서 탈퇴 동작이 발생하는 비율을 나타낸다고 하자. 그렇다면 가입 동작은 $(1-p)$의 비율로 발생하므로 이들이 발생하는 평균 키갱신 메시지 개수$A(p)$는 다음과 같다.

$$A(p) = (1 - 2p + pd) \lceil \log_d n \rceil + (1 - p).$$

우리는 [정리 3.4]로부터 $p = 1/2$인 경우 $d = 2$ 또는 $d = 3$인 경우에 $A(1/2)$가 최소값을 낳는다는 것을 증명하였었다. 그러나, 다음의 [정리 3.5]는 [정리 3.4]보다 더욱 강한 정리로서 만일 $p \geq 0.5$라면 $d = 2$ 또는 $d = 3$ 일 때 $A(p)$ 값이 최소가 된다는 것을 보여 준다. [부속정리 3.8]~[부속정리 3.10]은 [정리 3.5]를 유도하는 과정에 필요한 정리들이다.

[부속정리 3.8] $d \geq 4$및 $p \geq 1/2$라면 $1/2 \leq 1 - 2p + pd$이다.

(증명)

$$d/2 \leq 1 - 2p + pd$$
$$\rightarrow d \leq 2 - 4p + 2pd$$
$$\rightarrow 2(2p - 1) \leq d(2p - 1) \quad (3.2)$$

주어진 조건에서 $d \leq 4$이며 $(2p - 1) \leq 0$이므로 식 (3.2)는 항상 참이다. ■■■

주어진 n에 대하여 $A_d(p)$를 차수 d를 가지는 키트리(완전 정규 트리)의 $A(p)$라 하자. 그러면, $A_2(p) = \lceil \log_2 n \rceil + (1 - p)$가 된다. 다음 부속 정리는 $A_2(p)$와 $A_d(p)$$(d \geq 4)$의 관계를 설정해 준다.

[부속정리 3.9] $n(\geq 6)$ 및 $4 \leq d \leq n$을 만족하는 모든 정수 n 및 d에서 대하여 만일 $p \geq 1/2$라면 $d \geq 4$에 대하여 $A_2(p) \leq A_d(p)$이다.

(증명) [부속정리 3.6]으로부터 $n \geq 6$ 및 $4 \leq d \leq n$에 대하여 $\lceil \log_2 n \rceil \leq (d/2) \lceil \log_d n \rceil$ 임을 알 수 있다. 따라서, 다음이 성립한다.

$$A_2(p) = \lceil \log_2 n \rceil + (1 - p) \leq (d/2) \lceil \log_d n \rceil + (1 - p)$$

또한 [부속정리 3.8]로부터 $d \geq 4$및 $p \geq 1/2$라면 $(d/2) \leq 1 - 2p + pd$임을 알 수 있다. 따라서, $d \geq 4$에 대하여 다음이 성립한다.

$$A_2(p) \leq (1 - 2p + pd) \lceil \log_d n \rceil + (1 - p) = A_d(p). \quad ■■■$$

[부속정리 3.10] $n(\geq 6)$ 및 $4 \leq d \leq n$을 만족하는 모든 정수 n 및 d에서 대하여 만일 $p \geq 1/2$라면 $d = 2$ 또는 $d = 3$ 일 때 $A(p)$는 최소값을 갖는다.

(증명) [부속정리 3.9]로부터 $d \geq 4$에 대하여 $A_2(p) \leq A_d(p)$이므로 만일 $A_3(p) < A_2(p)$라면 $A_3(p)$가 $A(p)$의 최소

값이 되며 만일 $A_3(p) \geq A_2(p)$라면 $A_2(p)$가 $A(p)$의 최소값이다.

[정리 3.5] 완전 정규 트리 모델 T에서 탈퇴 동작의 발생율이 50% 이상이라면 키갱신 메시지를 최소화할 수 있는 구조는 T의 차수 d가 2 또는 3 인 경우이다.

[정리 3.5]은 $p \geq 0.5$인 경우 $A_2(p)$ 및 $A_3(p)$등 2개의 값들만을 비교함으로서 $A(p)$의 최소값을 구할 수 있다는 것을 보여 준다. 그러나, $p < 0.5$라면 모든 $A_d\,(2 \leq d \leq n)$ 값들과 비교를 하여야 할 것이다. 다음은 완전 정규 트리를 이용한 키트리 모델에서 키갱신 메시지 개수를 최소화할 수 있는 차수 d를 계산하는 알고리즘이다.

```
입력 : n 은 최대 사용자 수, p 는 탈퇴동작 발생율.
출력 : 최적화를 이룰 수 있는 그룹키의 차수 d.
if (p ≥ 0.5)
    if A2(p) < A3(p)
        d = 2 ;
    else
        d = 3 ;
    end if
else
    Ave = ∞ ;
    for (i = 2 ; i ≤ n ; i++)
        if (Ai(p) < Ave)
            d = i ;
            Ave = Ad(p) ;
        end if
    end for
end if
```

(알고리즘 3.1) 완전 정규 트리의 최적화 알고리즘

4. 가변트리를 이용한 그룹키 모델 최적화

본 절에서는 가변트리라는 새로운 트리 구조를 정의하여 트리를 기반으로 한 그룹키 관리 모델을 최적화할 수 있는 방안을 연구한다.

[정의 4.1] 가변 트리 T는 트리로서 모든 말단 정점들은 동일한 레벨에 위치하며 동일한 레벨에 위치한 내부 정점들은 동일한 개수의 자식 노드를 갖지만 서로 다른 레벨에 위치한 내부 정점들은 서로 다른 개수의 자식 노드를 가질 수 있다. 높이 h를 갖는 가변 트리 T의 차수 수열 $\mathrm{Deg}(T) = \{d_1, d_2, \cdots, d_{h-1}\}$는 T의 레벨 i, $1 \leq i \leq h-1$, 에서 각 내부 정점들이 갖는 자식 노드의 개수를 나타낸다. 차수 수열 $\mathrm{Deg}(T) = \{d_1, d_2, \cdots, d_{h-1}\}$에서 모든 $d_i (1 \leq i \leq h-1)$들이 2인 차수 수열을 "(2)-차수수열"이라 정의하며 모든 d_i $(1 \leq 1 \leq i \leq h-1)$들이 2 또는 3인 차수 수열을 "(2, 3)-차수수열"이라 정의한다. T의 높이는 키트리와 동일하게 정의한다.

완전 정규 트리는 모든 내부 정점들이 동일한 개수의 자식 정점을 갖는 구조적 단순함으로 인하여 최대 사용자 수에 따라서는 비 효율적인 데이터 구조가 될 수도 있다. 그룹키 서버가 키트리를 이용하여 관리하여야 할 최대 사용자 수가 n이라 하자. 만일 n값이 d^k의 형태의 정수라면 키트리에는 불필요한 말단 정점이 포함되지 않지만 만일 $n \neq d^k$라면 완전 정규 트리 에는 상당히 많은 수의 불필요한 말단정점 및 서브 그룹키들이 포함되게 된다. 예를 들어서, n = 100,000인 경우 다음 표는 서로 다른 트리 차수를 갖는 완전 정규 트리 들의 효율성을 보여 준다. <표 4.1.1>에서 사용자의 수를 100,000으로 채택한 것은 모든 그룹키 관리 메커니즘은 적어도 100,000명의 사용자를 효율적으로 수용할 수 있어야 한다는 SMuG 그룹의 권고 사항을 따른 것이다.

〈표 4.1.1〉 100,000명을 수용하기 위한 키트리의 효율성

차수	높이	JMsg(T)	LMsg(T)	SubKey(T)	최대 사용자 수
2	18	18	17	131,070	131,072
3	12	12	22	88,572	177,147
4	10	10	27	87,380	262,144
5	9	9	32	97,655	390,625
6	8	8	35	55,986	279,936

이때 만일 탈퇴 동작의 발생율이 30% (즉, $p = 0.3$)라면 3장의 (알고리즘 3.1)로부터 트리의 차수는 3일 때 최소한의 키갱신 메시지가 발생된다는 것을 알 수 있으며 <표 4.2>는 본 경우에서 완전 정규 트리의 효율성을 보여 준다. $\mathrm{ReKey}_T(p)$, $\mathrm{Enc}_T(p)$ 및 $\mathrm{SubKey}_T(p)$는 탈퇴 비율이 p일 때 발생하는 키갱신 메시지 개수, 암호화 개수 및 서브 그룹키 개수라 정의한다.

〈표 4.2〉 $p = 0.3$인 경우 최적화된 완전 정규 트리의 효율성

차수	높이	$\mathrm{Rekey}_T(p)$	$\mathrm{Enc}_T(p)$	$\mathrm{SubKey}_T(p)$
3	12	15	25	88,572

또한 <표 4.3>은 $p = 0.3$인 경우 $\mathrm{Deg}(T) = \{4^6 5^2\}$로 구성된 가변트리의 효율성을 보여 준다.

〈표 4.3〉 $p = 0.3$인 경우 최적화된 가변트리의 효율성

차수	높이	$\mathrm{Rekey}_T(p)$	$\mathrm{Enc}_T(p)$	$\mathrm{SubKey}_T(p)$
$\{4^6 5^2\}$	9	14.1	21.1	25,940

우리는 <표 4.2> 및 <표 4.3>을 비교함으로서 모든 효율성 파라메타에서 가변트리가 우월함을 알 수 있다. 특히

요구되는 서브 그룹키의 경우 완전 정규 트리는 가변트리에 비하여 약 4배 정도의 서브 그룹키를 필요로 한다는 것을 알 수 있다. 서브 그룹키들은 그룹키 서버에 의하여 모두 저장 및 관리되어야 하므로 최소한의 서브 그룹키를 사용하는 것은 그룹키 서버의 효율성 측면에서 매우 중요한 것이다.

가변트리가 완전 정규 트리에 비하여 진보된 효율성을 보이는 이유는 가변트리의 경우 그룹이 필요로 하는 최대 사용자 수에서 완전 정규 트리 보다 더욱 근접한 개수의 말단 정점들을 생성할 수 있기 때문이다. 예를 들어서, <표 4.2> 및 <표 4.3>에서 최대로 필요한 말단 정점의 개수가 100,000 일 때 완전 정규 트리의 경우 131,072개의 말단정점이 필요하지만 가변 트리의 경우 102,400로서 가변 트리가 필요한 말단 정점의 개수에서 완전 정규트리 보다 훨씬 근접하도록 구성할 수 있다는 것을 알 수 있다. 이것은 가변 트리가 트리 구조 측면에서 완전 정규 트리 보다 훨씬 유연하기 때문에 가능한 것이다.

가변트리를 이용한 그룹키 관리 구조의 효율성은 다음과 같이 계산할 수 있으며 증명과정은 단순한 것이므로 생략한다.

[정리 4.1] T를 $\mathrm{Deg}(T)=\{d_1, d_2, \cdots, d_{h-1}\}$의 차수를 갖는 가변트리라 하자. 그러면 T의 효율성은 다음과 같다.

(1) $\mathrm{JMsg}(T) = h$

(2) $\mathrm{LMsg}(T) = \sum d_i - (h-1)$

(3) $\mathrm{JEnc}(T) = 2(h-1)$

(4) $\mathrm{LEnc}(T) = \sum d_i - 1$

(5) $\mathrm{SKey}(T) = d_1 + d_1d_2 + \cdots + d_1 \cdots d_{h-3}\, d_{h-2}$ $(h \geq 3)$

가변트리 모델에서 가입 및 탈퇴동작이 각각 50%씩 균등하게 발생한다고 가정하자. 그렇다면 평균적으로 발생하는 키갱신 메시지 개수 A는 다음과 같다.

$$A = (1/2)(\mathrm{JMsg}(T) + \mathrm{LMsg}(T)) = (1/2)(\sum d_i + 1) \quad (4.1)$$

A는 $\sum d_i$가 최소가 될 때 최소가 되므로 가입 및 탈퇴 동작이 동일한 비율로 발생할 때 우리는 $\sum d_i$를 최소화 할 수 있는 수열 $\{d_1, d_2, \cdots, d_{h-1}\}$을 찾는 것이 목적이다. 그러나, 가변 트리는 항상 주어진 n 명의 사용자를 수용하여야 하므로 수열 $\{d_1, d_2, \cdots, d_{h-1}\}$로 구성되는 가변 트리의 말단 정점들의 개수 $d_1d_2 \cdots d_{h-1} = \Pi d_i$로서 이 값은 반드시 n보다 크거나 같아야 한다. 즉, 차수수열 $\{d_1, d_2, \cdots, d_{h-1}\}$은 $d_1d_2 \cdots d_{h-1} = \Pi d_i \geq n$의 조건을 만족시켜야만 한다.

[부속정리 4.1]~[부속정리 4.4] 및 [정리 4.2]은 가변트리 모델에서 가입 및 탈퇴 동작의 발생 비율이 50 : 50인 경우 (2, 3)-차수수열이 키갱신 메시지 개수를 최소화할 수 있는 구조임을 보여 준다.

[부속정리 4.1] 모든 $n(\geq 5)$에 대하여 $2n+1 < 2^{n-1}$이다.

(증명) 증명은 n에 대한 귀납법을 이용한다. $n = 5$ 이면 $2(5)+1 = 11 < 2^4 = 16$이다. 이제 $n(>6)$에 대하여 주어진 식이 성립한다고 가정하고 $2(n+1)+1 < 2^n$임을 증명한다.

$$2(n+1)+1 = 2n+1+2 < 2^{n-1}+2 < 2^{n-1}+2^{n-1} = 2(2^{n-1}) = 2^n$$

이며, 따라서 $2(n+1)+1 < 2^n$이다. ■■■

[부속정리 4.2] 모든 $n(\geq 7)$에 대하여 $s = d_1d_2 \cdots d_m \geq n$ 및 $p = d_1 + d_2 + \cdots + d_m \leq n$를 만족하는 (2)-차수수열 $\{d_1, d_2, \cdots, d_m\}$이 존재한다.

(증명) (1) $m = \lceil \log_2 n \rceil$이라 하자. 먼저 $s = n^{\lceil \log_2 n \rceil}$을 증명하자.

$$2^{\lceil \log_2 n \rceil} \geq n \rightarrow \lceil \log_2 n \rceil \geq \log_2 n \quad (4.2)$$

이며, 식 (4.2)는 $n \geq 1$에 대하여 참이므로 $s \geq n$이다.

(2) 다음은 $p = 2\lceil \log_2 n \rceil \leq n$임을 증명하도록 하자. 이것을 위하여 먼저 $2 \cdot \log_2 n < n-1$임을 $n(\geq 7)$에 대한 귀납법을 이용하여 증명하도록 한다. 그러나, $2 \cdot \log_2 n < n-1 \rightarrow \log_2 n^2 < \log_2 2^{n-1}$이므로 우리는 $n^2 < 2^{n-1}$을 n에 귀납법을 이용하여 증명하도록 한다. $n = 7$ 이면 $7^2 = 49 < 2^6 = 64$이다. 이제 주어진 식이 n에 대하여 성립한다고 가정하고 $(n+1)^2 < 2^n$임을 증명한다.

$$\begin{aligned}(n+1)^2 &= n^2 + 2n + 1 \\ &< 2^{n-1} + 2n + 1 \quad (\because \text{귀납 가정}) \\ &< 2^{n-1} + 2^{n-1} \quad (\because \text{[정리 4.2]}) \\ &= 2(2^{n-1}) = 2^n\end{aligned}$$

이며 따라서, $2\log_2 n < n-1$이다. 이제 [부속정리 3.1]에 의하여 $\lceil 2\log_2 n \rceil \leq \lceil n-1 \rceil$이다.

[부속정리 4.3] 모든 정수 $n(\geq 4)$에 대하여 $s = d_1d_2 \cdots d_m \geq n$ 및 $p = d_1 + d_2 + \cdots + d_m \leq n$를 만족하는 (2, 3)-차수수열 $\{d_1, d_2, \cdots, d_m\}$이 존재한다.

(증명) $n = 4, 5, 6$ 인 경우 다음과 같은 (2, 3)-차수수열들이 존재한다.

- $n = 4$ 인 경우 $\{2 \cdot 2\}$
- $n = 5$ 인 경우 $\{2 \cdot 3\}$
- $n = 6$ 인 경우 $\{2 \cdot 3\}$

$n \geq 7$인 경우 [부속 정리 4.2]은 n에 대하여 (2)-차수수열이 존재함을 보여 준다. (2)-차수수열은 (2, 3)-차수수열에 포함되므로 증명은 완료된다. ■■■

[부속정리 4.4] 가변 트리 T에서 가입 및 탈퇴 동작이 발생하는 비율을 50 : 50 이라 하고 키갱신 메시지의 개수를 최소로 하는 차수수열을 $s_{min}=\{d_1, d_2, \cdots, d_{k-1}\}$이라 하자. 그렇다면 s_{min} 중에는 (2, 3)-차수수열이 존재한다.

(증명) 가변 트리에서 주어진 사용자 수 n 에 대한 최적의 수열을 $s_{min}=\{d_1, d_2, \cdots, d_{k-1}\}$이라 하자. 만일 s_{min}이 (2, 3)-차수수열 이라면 증명은 완료된다. 만일 그렇지 않다면 $1 \le i \le h-1$에 대하여 $d_i \neq 2$ 및 $d_i \neq 3$인 d_i가 존재한다. 이제 s'_{min}을 s_{min}으로부터 d_i를 제거하고 d_i 대신에 부속 정리4.3에서 제시되는 차수수열 $\{d'_1, d'_2, \cdots, d'_t\}$ $(t \ge 2)$로 대체한 차수수열이라 하자 (s'_{min} 수열의 길이는 2 이상이 된다). [정리 4.4]로부터 $\sum_{i=1}^{t} d'_i \le d_i$ 및 $\prod_{i=1}^{t} d'_i \ge d_i$ 임을 알 수 있으므로 우리는 다음과 같은 부등식들을 얻는다.

$$\sum s'_{min} \le \sum s_{min}$$
$$\prod s'_{min} \ge \prod s_{min} \ge n$$

이제 만일 s'_{min}이 (2, 3)-차수수열 이라면 우리의 증명은 완료된다. 만일 그렇지 않다면 상기한 과정을 반복적으로 적용할 수 있다. ■■■

[부속정리 4.1]~[부속정리 4.4]로부터 우리는 다음 정리를 얻는다.

[정리 4.2] 가변 트리 T를 기반으로 한 그룹키 관리에서 가입 및 탈퇴 동작의 발생 비율이 50 : 50인 경우 키갱신 메시지를 최소화할 수 있는 T의 차수수열은 Deg(T)가 (2, 3)-수열차수인 경우이다.

우리는 키트리와 유사하게 가변 트리 모델에서 탈퇴동작의 발생 비율이 50%보다 큰 경우 키갱신 메시지 개수를 최소화할 수 있는 차수수열은 (2, 3)-차수수열임을 증명할 수 있다. 가변 트리 T를 기반으로 하는 그룹키 관리에서 탈퇴 동작이 발생하는 비율을 p라 하자. 따라서 Deg(T) = $\{d_1, d_2, \cdots, d_k\}$일 때 발생하는 평균적인 키갱신 메시지 개수를 $A(p)$라 하면 $A(p)$ 는 다음과 같다.

$$A(p) = h(1-2p) + p(\sum d_i + 1) \quad (4.3)$$

[정리 4.3] 가변 트리 T를 모델로 하는 그룹키 관리에서 탈퇴 동작 의 발생 빈도수 p가 0.5보다 크다면 키갱신 메시지 개수를 최소화하는 차수수열은 (2, 3)-차수수열이다.

(증명) 주어진 사용자 수 n 및 $p(> 0.5)$ 에 대하여 $s = \{d_1, d_2, \cdots, d_k\}$를 $A(p)$를 최소화하는 수열이라 하자. 만일 s가 (2, 3)-차수수열 이라면 증명은 완료된다. 만일 그렇지 않다면 $d_i \neq 2$ 및 $d_i \neq 3$인 d_i, $1 \le i \le h-1$, 가 s 내에 반드시 존재한다. s'를 s로부터 d_i를 제거하고 [부속정리 4.3]으로부터 정의된 수열로 대체한 차수수열 이라 하자. T 및 T' 를 수열 s 및 s'로부터 정의된 가변 트리 모델이라 하고 h 및 h'를 T 및 T'의 높이라 하자. 또한 $\sum d_i$ 및 $\sum d'_i$를 수열 s 및 s'의 합이라 하자. 그렇다면 명확히 $h < h'$이며 또한 [부속정리 4.3]으로부터 $\sum d_i \le \sum d'_i$이다. $A(p)$ 및 $A'(p)$를 탈퇴 동작의 발생 비율이 p 일 때 s 및 s'를 차수수열로 하는 가변 트리 모델의 평균 키갱신 메시지 개수라 하자. 그러면

$$A(p) - A'(p) = (1-2p)(h - h') + p(\sum d_i - d'_i)$$

가 된다. 그러나, $h - h' > 0$ 및 $\sum d_i \ge \sum d'_i$이고 또한 $0.5 < p \le 1$이면 $1 \le (1-2p) < 0$이므로 다음 관계가 성립한다.

$$A(p) - A'(p) = (1-2p)(h - h') + p(d_i - d'_i) > 0.$$

따라서, $A(p) > A'(p)$임을 알 수 있다. 만일 s'가 (2, 3)-차수수열이라면 증명은 완료되며 만일 (2, 3)-차수수열이 아니라면 상기한 과정을 반복적으로 적용할 수 있다. ■■■

[정리 4.2] 및 [정리 4.3]에서 언급되는 (2, 3)-차수수열은 실제로는 유일한 것이 아니므로 우리는 가능한 모든 (2, 3)-차수수열 들을 모두 고려하여야 한다. [정리 4.4]은 이미 널리 알려진 정리이며 우리는 이를 이용하여 [정리 4.5]에서 최대 n명의 사용자를 포함하는 가변 트리 T에서 나타날 수 있는 모든 (2, 3)-수열의 최대 개수를 분석한다.

[정리 4.4] X가 t개의 원소를 가진 집합이라고 할 때, 반복이 허용된다면, X에서 순서에 관계없이 k개의 원소를 선택하는 방법의 수는 다음과 같다.

$$C(k+t-1, t-1) = C(k+t-1, k)$$

[정리 4.5] $n (\ge 4)$개의 말단 정점을 포함하는 가변 트리 T의 (2, 3)-차수수열의 길이를 k라 하자. 그렇다면 k의 범위는 $2 \le k \le \lceil \log_2 n \rceil$ 이며 T를 구성할 수 있는 (2, 3)-차수수열의 총 개수는 $\sum_{k=0}^{\lceil \log_2 n \rceil} C(k+1, 1)$이다.

(증명) $n \ge 4$이므로 (2, 3)-차수수열로 구성된 가변 트리의 높이는 반드시 3보다 크거나 같아야 한다. 또한 n개의 말단 정점을 포함할 수 있는 최소한의 높이를 가지는 완전 이진 정규 트리의 높이는 $\lceil \log_2 n \rceil + 1$이다. 따라서, k의 범위는 $2 \le k \le \lceil \log_2 n \rceil$ 이다.

각 $k (2 \le k \le \lceil \log_2 n \rceil)$에 대하여 서로 다른 (2, 3)-차수수열은 [정리 4.8]에 의하여 모두 $C(k+2-1, 1) = C(k+1, 1)$ 이므로 가능한 (2, 3)-차수수열의 총 개수는 $C(k+1, 1)$이 된다. ■■■

탈퇴 동작이 50% 이하인 경우에는 우리는 가능한 모든 가변트리 들을 고려하여야 하며 따라서 (2, 3)-차수수열 이외의 모든 차수수열 들을 고려하여야만 한다. 높이 h 및 n개의 말단 정점을 갖는 가변트리 T의 차수수열을 $\{d_1, d_2, \cdots, d_{h-1}\}$이라 하자. 만일 우리가 $d_i\,(1 \le i \le h-1)$에 특정한 제약을 부여하지 않는다면 개개 d_i들은 2부터 n까지 조사가 되어야 하므로 만일 n값이 충분히 크다면 이것은 현실적으로 조사 가능한 범위를 벗어나게 된다. 따라서, 우리는 조사되어야 할 수열의 개수 및 각 수열의 d_i 값들을 최소화하여야 한다. 식 (4.3)으로부터 p및 h가 상수 값인 경우 $A(p)$의 최소값은 d_i에 의하여 결정된다는 것을 알 수 있다. 따라서, 동일한 높이를 가지는 가변 트리들 사이에서 각 가변 트리 들의 차수수열의 합이 최소인 가변 트리가 동일한 높이를 갖는 가변트리 들 사이에서 $A(p)$ 값을 최소로 한다는 것을 알 수 있다. [정리 4.6]은 알고리즘이 고려하여야 할 가변 트리의 높이의 상한선을 제시한다.

[정리 4.6] 가변 트리 T의 차수수열 $\{d_1, d_2, \cdots, d_k\}$ $(d_i \ge 2,\ 1 \le i \le k)$가 $\lceil \log_2 n \rceil$의 길이를 갖는 (2)-차수수열과 동일한 길이를 갖으며 (즉, $k = \lceil \log_2 n \rceil$) $d_1 \cdot d_2 \cdots d_k \ge n$을 만족하지만 (2)-차수수열과는 다른 수열이라 하자. 그렇다면 정수 $n\,(\ge 2)$에 대하여 $d_i > 2$이다.

(증명) $\{d_1, d_2, \cdots, d_k\}$는 2-수열과는 다르므로 최소한 한 개의 $d_j\,(1 \le j \le k)$는 2보다 크다. 따라서, $d_j > 2$라 하자. 그렇다면, $d_i\,(1 \le i \le k-1,\ i \ne j)$의 크기는 최소한 2이므로 $\sum d_i$의 최소값은 다음과 같다.

$$\begin{aligned}\sum d_i &= d_1 + d_2 + \cdots + d_k \\ &\ge 2(\lceil \log_2 n \rceil - 1) + d_j = 2\lceil \log_2 n \rceil - 2 + d_j\end{aligned}$$

따라서,

$$\begin{aligned}\sum d_i - 2\lceil \log_2 n \rceil &\ge ((2\lceil \log_2 n \rceil - 2 + d_j) - 2\lceil \log_2 n \rceil \\ &= d_j - 2 > 0 \qquad (\because\ d_j > 2)\end{aligned}$$

이므로 $\sum d_i > 2\lceil \log_2 n \rceil$ 이다. ■■■

사용자 수 $n\,(\ge 2)$을 말단 정점으로 포함할 수 있는 최소한의 높이를 갖는 (2)-차수수열로 구성된 가변 트리를 T_{bin}라 하자. 그렇다면 T_{bin}의 높이 $h = \lceil \log_2 n \rceil + 1$, $\mathrm{Deg}(T_{bin}) = \{2^{\lceil \log_2 n \rceil}\}$이며 T_{bin}의 차수수열의 합은 2이다. [정리 4.6]은 T_{bin}과 동일한 높이를 가지는 다른 가변 트리 들의 차수수열의 합은 2보다 크다는 것을 증명하고 있으며 따라서 우리는 T_{bin}과 동일한 높이를 갖는 가변 트리 들 중 완전 이진 정규 트리 만을 고려하는 것으로 충분하다는 것을 보여 준다. 또한 모든 가변 트리 들의 높이는 최소한 2 이상이어야 하므로 따라서 알고리즘에서 고려하여야 할 가변 트리 들의 높이의 범위는 $2 \sim \lceil \log_2 n \rceil$ 이다.

$n\,(\ge 7)$개의 말단 정점을 갖는 가변 트리 T의 차수수열 $\mathrm{Deg}(T)$를 $\{d_1, d_2, \cdots, d_k\}$라 하자. 다음 정리는 주어진 차수수열의 길이 k에 대하여 알고리즘이 고려하여야할 d_i의 범위를 밝혀 준다.

[정리 4.7] n개의 말단정점을 갖는 가변 트리 T의 차수수열을 $\{d_1, d_2, \cdots, d_k\}$라 하자. 만일 $d_j > 2\lceil \log_2 n \rceil - 2(k-1)$ $(1 \le j \le k)$이면 $\sum d_i > 2\lceil \log_2 n \rceil$이다.

(증명) $d_i\,(1 \le i \le k)$는 2보다 크거나 같기 때문에 d_i의 하한 값은 다음과 같다.

$$\begin{aligned}\sum d_i &\ge 2(k-1) + d_j \\ &> 2(k-1) + 2\lceil \log_2 n \rceil - 2(k-1) \\ &= 2\lceil \log_2 n \rceil .\end{aligned}$$ ■■■

[정리 4.7]에 의하여 n개의 말단정점을 갖는 가변트리 T (차수수열은 $\{d_1, d_2, \cdots, d_k\}$)의 각 높이에서 고려하여야 할 $d_i\,(1 \le i \le k)$의 범위는 $2 \le d_i \le 2\lceil \log_2 n \rceil - 2(k-1)$임을 알 수 있다.

다음은 상기한 정리들을 기반으로 하여 최적의 가변 트리 모델을 계산하는 알고리즘 이다. 본 알고리즘에는 $p \le 0.5$인 경우 즉, (2, 3)-차수수열을 포함한 모든 차수수열을 고려하는 경우만을 보였다. $p > 0.5$인 경우는 (2, 3)-차수수열만을 고려하는 것으로서 본 알고리즘과 매우 유사하므로 생략한다. (알고리즘 4.1)에서 수열 $s = \{s_0, s_1, \cdots, s_{t-2}\}$ 및 $b = \{b_0, b_1, \cdots, b_{t-1}\}$ $(t = \lceil \log_2 n \rceil)$들은 $s_i = 2\ (0 \le i \le t-2)$ 및 $b_j = 2\ (0 \le j \le t-1)$로 초기화되었다고 가정한다. 또한 $r_i \leftarrow x\ [u \le i \le v]$는 부분수열 $r_u \sim r_v$들의 값을 x로 대체하는 동작이라고 정의한다.

```
입력 : n은 최대 사용자수, p는 탈퇴 동작 발생율, 수열 s 및 b.
출력 : 최적화를 이룰 수 있는 가변 트리 차수수열 b = {b1, b2, ···,
       bm-1} ( m 값은 알고리즘 참조)

Opt = A2(p)
stop = 4
m = ⌈log2n⌉ - 1

for (k = ⌈log2n⌉ -1 ; k ≥ 2 ; k--)
    h = k + 1
    for (i = 0 ; i < C(2⌈log2n⌉ -k, k) ; i++)
        A = h(1-2p) + p( Σ_{j=0}^{m-2} Sj + 1)

        if ( Π_{j=0}^{m-2} sj ≥ n AND A < Opt)
            Opt = A
            bj ← sj[0 ≤ j ≤ m-1]
        end if

        if sm-1 ≠ stop
            sm-1 = sm-1 + 1
        else
            pivot = 0
            for ( c = m - 1 ; c 0 ; c--)
                if (sk ≠ stop)
                    pivot = sk
                    break ;
```

```
            end for
        end if

        if (k ≥ 0)
            s_j ← pivot + 1 [k ≤ j < m]
    end for

    m = m - 1
    stop = stop + 2
    s_i ← 2 [0 ≤ j < m]
end for
```

(알고리즘 4.1) 가변 트리의 최적화 알고리즘

[정리 4.8] (알고리즘 4.1)에서 고려하는 총 차수수열의 개수는 다음과 같다.

$$\sum_{k=2}^{\lceil \log_2 n \rceil - 1} C(2\lceil \log_2 n \rceil - k, k)$$

(증명) (알고리즘 4.1)에 의하여 고려되고 있는 수열의 길이를 k라 하자. 이 경우 모두 $2\lceil \log_2 n \rceil - 2(k-1) - 2 + 1 = 2\lceil \log_2 n \rceil - 2k + 1$개의 서로 다른 d_i가 고려되므로 [정리 4.4]에 의하여 고려되는 차수수열의 개수는 $C(2\lceil \log_2 n \rceil - 2k + 1 + k - 1, k) = C(2\lceil \log_2 n \rceil - k, k)$이며 k의 범위는 $2 \le k \le \lceil \log_2 n \rceil - 1$ 이므로 총 차수수열의 개수는 $\sum_{k=2}^{\lceil \log_2 n \rceil - 1} C(2\lceil \log_2 n \rceil - k, k)$이다. ■

상기한 알고리즘은 n = 100,000 및 서로 다른 p값에 대하여 다음 표와 같은 효율성을 보인다.

〈표 4.4〉 n = 100,000인 경우 가변트리의 효율성

p	Deg(T)	h	ReKey$_T$(p)	SubKey$_T$(p)	Enc$_T$(p)
0.1	$\{6, 7^6\}$	7	9.8	16,806	14.8
0.2	$\{4, 5^5, 8\}$	8	12.4	15,624	18.4
0.3	$\{4^6, 5^2\}$	9	14.1	25,940	21.1
0.4	$\{3^8, 4^2\}$	11	15.4	36,084	24.4
0.5	$\{3^8, 4^2\}$	11	16.5	36,084	25.5
0.6	$\{2^4, 3^8\}$	13	17.2	52,494	28.2
0.7	$\{2^{16}\}$	17	17.3	131,070	33.3
0.8	$\{2^{16}\}$	17	17.2	131,070	33.2
0.9	$\{2^{16}\}$	17	17.2	131,070	33.1

5. 결 론

본 연구의 주 목적은 멀티캐스트 통신에 정보보호 기능을 제공하기 위하여 필수적으로 요구되는 그룹키 관리 메커니즘을 연구하는 것이다.

그룹키 관리에 트리 모델을 적용하는 것은 [1]및 [2]에 의하여 최초로 제안되었지만 여기서 제시된 트리 모델은 완전 정규 트리로서 사용자 수에 알맞은 크기의 키트리 모델을 생성하는 것이 어려울 수가 있다. 따라서, 본 연구에서는 완전 정규 트리 보다 구조적인 면에서 더욱 유연성을 갖는 트리 모델을 정의 및 분석하여 키트리 모델의 최적화를 시도하였다.

이러한 목적을 위하여 2절에서는 기존에 [1]에서 제시된 키트리 모델 및 관리 알고리즘들을 분석하였다. 분석된 내용들을 바탕으로 3절에서는 사용자의 탈퇴 동작을 위해서는 키트리는 트리의 차수가 2인 경우에 최적의 구조가 되며 가입 동작을 위해서는 트리의 차수가 n인 경우에 최적의 구조가 된다는 것을 증명하였다. 이와 같이 사용자의 가입 및 탈퇴 동작에 대하여 서로 상반된 트리 구조에서 최적의 구조가 된다는 것은 우리가 사용자의 가입 및 탈퇴 비율을 감안한 상태에서 키트리 모델을 생성하여야 한다는 것을 의미하게 된다는 것을 제시하였다. 사용자의 가입/탈퇴 비율을 고려하였을 때 완전 정규 트리 모델의 경우 탈퇴비율이 50% 이상인 경우 트리의 차수가 2 또는 3인 경우에 최적의 키트리 모델이 된다는 것이 또한 증명되었다. 완전 정규 트리 모델은 구조상의 경직성으로 인하여 사용자 수가 d^k 형태가 아닌 경우에는 사용자의 수에 알맞은 크기의 키트리를 생성할 수 없기 때문에 4절에서는 가변 트리라는 새로운 트리 구조를 정의 및 분석하였다. 가변 트리는 각 레벨에서의 자식 정점의 개수는 동일하지만 서로 다른 레벨에서는 서로 다른 개수의 자식 정점들을 가질 수 있으므로 완전 정규 트리에 비하여 더욱 유연한 구조를 가지게 되며 따라서 사용자 수에 따라 알맞은 크기의 키트리를 생성할 수 있다. 또한 가변트리에서도 완전 정규 트리 모델과 유사하게 사용자의 탈퇴 비율이 50%를 넘는 경우 (2, 3)-차수수열을 갖는 가변 트리 들이 최적이 된다는 것이 증명되었다. 또한 3절 및 4절에서는 완전 정규 트리 모델 및 가변트리 모델들을 대상으로 모든 가입/탈퇴 비율에서 최적의 키트리 모델을 생성할 수 있는 알고리즘들이 각각 제안되었다.

그룹키 관리에 관한 연구는 멀티캐스트의 활용이 점차적으로 현실화 되어감에 따라 더욱 발전될 것으로 예상된다. 본 연구에서는 비록 완전 정규 트리 보다는 유연한 가변트리 모델을 새로이 제시하였지만 향후에는 가변트리 보다 더욱 유연한 구조를 가진 트리 모델에 대한 연구가 진행될 것으로 기대된다.

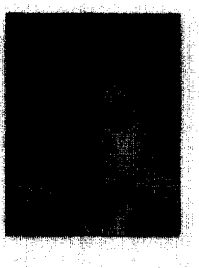
참 고 문 헌

[1] Chung Kei Wong, Mohamed Gouda, Simon S. Lam, "Secure Group Communications Using Key Graphs," ACM SIGCOMM '98, 1998.

[2] Debby M. Wallner, Eric J. Harder, "Key Management for Multicast : Issues and Architecture," internet draft, draft-wallner-key-arch-01.txt, 1998.

[3] M. J. Moyer, J. R. Rao, P. Rohatgi, "Maintaining Balanced Key Trees for Secure Multicast," internet draft, draft-irtf-smug-key-tree-balance-00.txt, 1999.

[4] Thomas Hardjono, Brad Cain, Indermohan Monga, "Intra-domain Group Key Management Protocol," internet-draft, draft-irtf-smug-intragkm-00.txt, 2000.
[5] Mitra S., "Iolus : A Framework for Scalable Secure Multicast," ACM SIGCOMM'97, 1977.
[6] Balenson, "Key Management for Large Dynamic Groups : One-Way Function Trees and Amotized Initialization," internet-draft, draft-balenson-groupkeymanagement-oft-00.txt.
[7] D. Harkins, D. Carrel, "The Internet Key Exchange (IKE)," RFC2409, 1998.
[8] Harrary, "Graph Theory," Addison-Wesley Publishing Co., 1972.



한 근 희

e-mail : kehan@kongju.ac.kr
1986년 건국대학교 물리학과 졸업(학사)
1992년 Univ. of Central Oklahoma 응용수학과졸업(이학석사)
1996년 Univ. of Oklahoma 컴퓨터과학과 졸업(이학박사)
1996년~2000년 한국전자통신연구원 재직(팀장)
1999년~2000년 미국 NIST 객원연구원
2000년~현재 공주대학교 응용수학과 조교수
관심분야 : 알고리즘, 그래프, 정보보호, 멀티캐스트