CHAPTER 4

BATCH SIZE OPTIMIZATION FOR IPTV

Secure group communication is an important design issue for profitable applications like IPTV. The security of the group is guaranteed by the process of rekeying and it can be immediate rekeying or batch rekeying. Selection of the rekeying strategy is mainly depends on the type of application. For large dynamic group immediate rekeying may introduce considerable overhead in communication and also may not be scalable. The overhead is imposed by the constraints on the bandwidth and also authentication of the rekey messages. To cope with this batch rekeying is proposed and it amasses the join and leave request of the subscriber and performs rekeying based on interval. Forward and backward secrecy are not strictly satisfied and relaxed for a while (Joe Prathap & Vasudevan 2008).

Setia et al (2000) described an approach for scalable group rekeying for secure multicast using periodic group rekeying, called Kronos. The inefficiency of individual rekeying under dynamic and large networks and compared the performance of Kronos with other key management protocols using simulation. Li et al (2001) designed a batch-rekeying algorithm, called keygem, to improve scalability and performance of a large and dynamic group. Recently, Lazos & Poovendran (2003) proposed the use of location-aware batch rekeying of key hierarchies in wireless ad hoc networks. However, earlier research in identification of optimal rekeying parameters for batch rekeying is minimal. Batch rekeying mainly depends on batch size and interval. The second phase of the research work focuses on the
identification of the optimal batch size for the batch rekeying interval. Identification of optimal batch size has an advantage of improving the performance and defining the period of security violation very clearly.

4.1 **BATCH REKEYING PARAMETERS**

Batch rekeying is an efficient mechanism of broadcasting multimedia contents in a secured manner. Though the security is violated for short span of time the communication overhead imposed by this approach is considerably less (Wee hock Desmond Ng & Zhilli Sun 2005). The bandwidth consumed for batch rekeying is less, which will improve the channel broadcasting potential noticeably. The major focus in batch rekeying is identification of optimal rekeying interval, which should reduce the waiting period of the subscribers as well as an efficient strategy to reduce rekeying overhead by trading off a secrecy violation. Batch rekeying is performed in two different steps as shown in the Figure 4.1. The first step is called as rekey request collection and the next one is called as new session key generation.

1. **Rekey request collection**: The rekeying request is raised due to subscriber’s join or leave. In this phase all the rekeying request are collected and the subscriber’s group is assorted into active subscribers and passive subscribers. Active subscribers group encompasses the members whose subscription is alive. Passive subscribers group comprises of the members who instigated the leave request or expelled by the service provider due to expiry of subscription.

2. **New session key generation**: The new session key is computed and communicated only to the existing subscribers.
In batch rekeying, the rekeying is performed within regular time slots called as rekeying intervals $RI$. During RI the various rekeying request are collected. The time taken for rekeying is denoted as batch rekeying interval $BRI_k$ where $k$ denoted the $k^{th}$ batch. If the rekeying request is not during the rekeying interval $RI$, the join as well as leave rekeying request waits for a period known as delayed batch join and delayed batch leave. The pictorial representation of batch rekeying intervals is illustrated in Figure 4.2.

The parameters that determines the performance of batch rekeying are

1. $K$- Total number of keys needed to process the rekeying request for every batch.
2. $E$- Total number of encryptions needed per batch.
3. $C_k$ – Average the cost of generating and disseminating the new session key.
4. $C_e$ - Average the cost of encryption.
The batch rekeying time BRI,

\[
BRI \geq \sum_{i=0}^{i=n}(K \cdot C_i) + \sum_{i=0}^{i=n}(E \cdot C_i)
\]  

Optimization of batch rekeying can be by differentiating the rekeying request into trusted and untrusted requests. Depends on the application the batch size import two different constraints are

1. Probability of security violation
2. Delay of rekeying

However, it is marked as two different constraints influenced strongly by each other.

The rekey server considers join as well as leave request as rekeying request. Since the subscriber joins the group by completing the admission process managed by resource admission control, the join request is assumed as trusted request. The new subscriber cannot become a malicious subscriber immediately. However, the leave can be either trusted leave or untrusted leave. The group controller upon the expiration of subscription can initiate the leave request characterized as trusted leave. The subscriber may use the services of IPTV and attack the setup called as malicious subscriber and to expel from the group instantly characterized as untrusted leave.

The nature of batch rekeying is to accumulate the rekeying request up to the threshold value and run the rekey algorithm (Jau chuan ke et al 2006). Most of the security challenges are thrown to this batch rekeying system. The vulnerability period is the period where the security is violated and identified from the probability of secrecy violation, which in turn
measures the trustworthiness of the system. The level of trustworthiness determines the size of the batch.

4.1.1 Trustworthiness of a System

Though IPTV is a paid service, if the security is violated for a small amount of time the loss at the broadcast server is minimal when compared to the overhead induced by the immediate rekeying. Depends on time, the trustworthiness of the system varies (Hyeokchan Kwon et al 2009). Generally, the trustworthiness of a system is the ratio between total request and untrusted request. Since the join follows a detailed resource admission control mechanism assumed as trusted, only leaves are classified as trusted and untrusted rekeying request. The probability of trustworthiness is defined as

\[
P_t = \frac{\text{Number of trusted leave rekeying request}}{\text{Total number of rekeying request}} \quad (4.2)
\]

4.1.2 Secrecy Violation

The crypto-period linked with a session key is governed by numerous application-specific considerations. The significance of the data is to be examined first, and level of data compromise should be defined.

The period of secrecy violation is the period where the forward secrecy is violated due to leave, particularly untrusted leave. The probability of secrecy violation is specified as \(P_{sv}\).

\[
P_{sv} = \frac{\text{Period of forward secrecy violation}}{\text{Total batch processing time}} \quad (4.3)
\]
Assume the arrival rate is uniform i.e. \( \lambda = 1 \) and the service rate \( \mu = 0.5 \). Generally, any network consists of communication delay. However, assume under ideal conditions the communication delay may be one. The batch rekey time and the threshold alone contributes to the calculation of secrecy violation. The Figure 4.3 shows the relationship between batch size and secrecy violation probability.

![Figure 4.3 Batch size vs. probability of secrecy violation](image)

**Figure 4.3 Batch size vs. probability of secrecy violation**

The batch size is to be determined by evaluating the probability of secrecy violation and delay in rekeying. The delay plays a major role in determining the batch size. Delay is application specific.

### 4.2 DELAYS IN BATCH REKEYING

In IPTV, the subscriber joins to any particular channel or a bundle of channel by completing the registration and authorization procedure. But the subscriber cannot view the programs immediately after the join. Whenever the subscriber leaves from the channel, the services offered is not stopped immediately. However, application of batch rekeying is an ideal solution for IPTV in terms of communication cost; it instigates the different delays in the group (Shoufan & Huss 2009). In order to improve the performance of
rekeying the delays in batch rekeying is to be addressed properly. Delays in batch rekeying can be categorized as

1. Delayed Batch Join (DBJ) and
2. Delayed Batch Leave (DBL).

4.2.1 Delayed Batch Join

DBJ is defined as the difference between the joining request initiated by the subscriber and joining approval by distributing the new session key. The delay depends on the arrival of a subscriber. Assume that the subscriber I initiates a join request at time $T$ where the batch rekeying is in progress for $k^{th}$ batch. The next batch starts after collecting the entire rekey request for $k+1^{th}$ batch. Since the subscriber I instigated the join request at the beginning of $k^{th}$ batch, the delay incurred is maximum as shown in Figure 4.4.

![Figure 4.4](image-url)

**Figure 4.4 Batch join at the start of an interval**

In a different scenario the user $j$ initiates a join request at the end of the interval $RI$. The request of $j^{th}$ subscriber will be processed with a minimal delay as shown in Figure 4.5.
Figure 4.5 Batch join at the end of an interval

Delay in batch join can be calculated for the different cases using the formula 4.4

\[
DBJ = T - (RI \lor BRT_{prev}) + BRT_{next}
\]  \hfill (4.4)

Where  \( BRT_{prev} \) denotes the batch rekeying time before or at the time \( T \)

\( BRT_{next} \) denotes the batch rekeying time after \( T \)

4.2.2 Delayed Batch Leave

Delayed batch leave is delineated as the difference between the leave request initiated by the subscriber and endorsed by disseminating the new session key. The delay depends on the arrival of leave request and the current batch processing time. Assume that the subscriber \( I \) initiates a leave request at time \( T \) where the batch rekeying is in progress for \( k^{th} \) batch. The next batch starts after collecting the entire rekey request for \( k+1^{th} \) batch. Since the subscriber \( I \) instigated the leave request at the beginning of \( k^{th} \) batch, the delay incurred is maximum as portrayed in Figure 4.6.
In another situation, the subscriber $j$ initiates a leave request at the end of the interval $RI$. The request of $j^{th}$ subscriber will be processed with a minimal delay as illustrated in Figure 4.7.

Delay in batch join can be calculated for the different cases using the equation 4.5

$$DBL = T \left( RI \lor BRT_{prev} \right) + BRT_{next}$$  \hspace{1cm} (4.5)

Where $BRT_{prev}$ denotes the batch rekeying time before or at the time $T$

$BRT_{next}$ denotes the batch rekeying time after $T$
4.3 ARRIVAL RATE CALCULATION

Subscribers join and leave initiates rekeying requests called as arrival rate, which is the key parameter to the queuing system (Laxmi & Gupta 1999). The arrival rate is the major influencing parameter in the performance of batch rekeying for IPTV system.

Figure 4.8 shows the inter arrival rate during the period 9.00 pm to 10.00 pm. This shows that the subscribers' arrival is a pure random process. In order to analyze the subscribers' join and leave frequency mathematically, the interarrival rate and the amount of resources in use are to be identified. This can be expressed as a special class of integrands known as stochastic integral.

![Inter-Arrival Time](image)

**Figure 4.8 Inter arrival time of subscribers during 9.00 10.00 pm**

The stochastic process \( A = \{ A(t) \mid -\infty < t < \infty \} \) is nonhomogeneous Poisson with mean rate function \( \lambda \), if it has independent Poisson increments (Samuel Blanquart & Nicolas Lartillot 2006). It means that for all \( s < t \), \( \int_s^t \lambda(t)dt \), integrable locally.
Let \( \{S_n \mid n = 1, 2, \ldots\} \) be an independent and identically distributed sequence of random variables and a function \( \varphi : \mathbb{R}^2 \to \mathbb{R} \) to be an \textit{integrand} only if it is a non-negative measurable function. Then the stochastic integral with respect to a nonhomogeneous Poisson process \( A \) is

\[
\int \varphi(S_{A(t)}, t) \, dA(t) \equiv \sum_{n=1}^{A(t)} \varphi(S_{n}, A_n)
\]  

(4.6)

where \( dA(t) \) is \( A(t) - A(t^-) \)

\( A_n \) is the time of the \( n \)th arrival in the interval.

Unlike stochastic integration over Brownian motion (Philip 2004), this can be defined as a sample path integration.

The arrival rate at the time \( t \) for the rekeying server is \( \lambda(t) \). The service rate is \( \mu(t) \). Together the arbitrary arrival rate and mean service rate determines the load at the time \( t \). Let \( L(t) \) be the load at the time \( t \).

\[
L(t) = \lambda(t) * \mu(t)
\]  

(4.7)

Since the arrival follows Poisson distribution and the service is exponential distribution the Markovian principle can be applied to calculate the steady state probability. Let the steady state probability be \( Q \).

Let \( Q(t) \) is Number of susbcribers in progress at time \( t \).

\( D(t) \) is Delay in service at time \( t \).

Using the stochastic integration the value of \( Q(t) \) and \( D(t) \) can be calculated as

\[
Q(t) = \int_{-\infty}^{t} (S_{A(t)}, t) \, dA(t)
\]  

(4.8)
\[ D(t) = \int_{-\infty}^{t} (S_{A(t)}, t) \, dA(t) \]  
\[ (4.9) \]

Where \( S_{A(t)} \) = sequence of random Poisson variable

\[ dA(t) \text{ is } A(t) - A(t^-) \]

Prediction of the system load is the key ingredient in planning for a queue based system. (Lawrence Brown et al 2005). Statistically, this prediction is based on a combination of the observed arrival time to the service time.

To predict the load of a queuing system, the confidence interval levels of the system for arrival and service is to be identified. In order to identify the confidence interval level the statistical property, Coefficient of Variation (CV) is used. For any non negative variable \( w \) with positive mean and variance, the CV is defined as

\[ CV(w) = \frac{\text{Standard Deviation}(w)}{\text{Mean}(w)} \]  
\[ (4.10) \]

In this model, the arrival is predicted for the near future by summing up all the arrivals of present and immediate past. Such predictions should be valid for future on which the arrival behavior follows the same pattern as those for that period of data. Assume the arrivals are considered for \( n \) days. The total number of arrivals are \( N_{jk} \), where \( j \) indicates the time of the day and \( k \) indicates the number of days.

Since the arrival to the rekeying server is assumed as Poisson process

\[ N_{jk} = \text{Poisson}(\lambda_{jk}) \]

\[ \hat{\lambda}_{jk} = R_i \times \tau_k^{-1} \varepsilon_k \]  
\[ (4.11) \]
Where \( R_i \) is the random effect during day \( i \)
\( \tau_k \) is the fixed hourly effects during hour \( k \).
\( \varepsilon_k \) is the random error in the prediction.

The expected number of subscribers during the interval \( k \) of a day

\[ \text{does not depend on the subscribers during the complete day. According to} \]

\[ \text{the theory of probability,} \]

\[ \sum \tau_k = 1 \quad (4.12) \]

Using the notion of stochastic process, user activity is defined as

the number of active subscribers at any time \( T \). The total number of IPTV

subscribers observed are around 350. If the average subscribers of the system

during the continuous time interval \( T \) is known with the help of the stochastic

process the average can be measured easily.

Figure 4.9 shows the average number of active subscribers during

the day where the diurnal trend is clearly seen. The user activity peaks in the

evening starting at approximately 07 pm and ending at approximately 08 pm.

The weekends have more active users between 8 am to 6 pm but less in the

evening compared to weekdays. The average peak number of active hosts is

around 240. It is necessary to see how the number of active IPTV users

evolve over time as this has important relations and is important to determine

the arrival rate for the rekey system.
Figure 4.9 Average number of active subscribers

The rekey server in IPTV runs the rekeying algorithm according to the batch rekeying interval. Batch size plays a vital role in identifying the optimal rekey interval. The threshold of a batch should not be high with respect to secrecy violation. The probability of secrecy violation reduces if the delay is minimal. Since the rkey time and batch, employs the principles of queuing theory, the rate of rekeying request can be predicted. If the rekey time is adjusted based on the peak and non peak periods of subscribers by applying the principles of queuing theory the load to the queue server can be predicted.

4.4 IDENTIFICATION OF OPTIMAL BATCH SIZE

Understanding of user activities is essential to many system design and engineering tasks such as evaluation of various design options, optimal system parameter tuning, improving customer care, and defining effective system care procedures to minimize service impact. User activity plays a
major role, in identifying the optimal batch size. In IPTV, the batch size mainly depends on the two factors

1. Arrival rate
2. Delay

The delay and the arrival rate of a system is analyzed throughly using the derived equations. Based on these parameters the total time spent by the subscribers of IPTV categorized into three sessions.

1. On –Sessions
2. Off-Sessions
3. Channel Zapping Session

The subscriber of IPTV goes into on and off sessions during the subscription period. An on session is the period where the subscriber switches on and enjoys the service and goes to off mode. Similarly an off session is the time period when the subscriber switched off and again switches on. The Subscription period is the combination of on and off session. Since two random parameters are used a mixture-exponential distribution can be applied to model the subscription period.

The probability density function of a mixture – exponential distribution is

$$f(x) = \sum_{i=1}^{n} a_i \cdot \lambda_i e^{-\lambda_ix}$$

(4.13)

Where the mean of the exponential distribution is $1/\lambda_i$ and the mixture parameter is $a_i$ and $\sum_{i=1}^{n} a_i$. Because the subscriber watches either at night or in the morning the distribution is skewed. The model built helps to analyze the skewness of the distribution (Marcia D Branco & Dipak K Dey 2001). The resultant distribution function is heavy tailed distribution (Anja Feldmann & Ward Whitt 1997)
To fit the distribution model the likelihood and expectation maximization are applied (Yair Weiss 1998). The random parameter $\lambda$ is the arrival rate of subscribers into the session and $\alpha$ is skewness ranging maximum of one. However, the large data set can be generated only maximum likelihood estimate parameters are taken into consideration. The distribution is iterated up to $n=3$.

Table 4.1 Characteristics of sessions in IPTV

<table>
<thead>
<tr>
<th>Sessions</th>
<th>$\lambda_1$</th>
<th>$a_1$</th>
<th>$\lambda_2$</th>
<th>$a_2$</th>
<th>$\lambda_3$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-Session</td>
<td>1.3e-2</td>
<td>0.3</td>
<td>3.3e-3</td>
<td>0.66</td>
<td>2.3e-4</td>
<td>0.04</td>
</tr>
<tr>
<td>Off-Session</td>
<td>3.2e-2</td>
<td>0.19</td>
<td>2.5e-3</td>
<td>0.75</td>
<td>2.4e-4</td>
<td>0.06</td>
</tr>
<tr>
<td>Channel Session</td>
<td>2.1</td>
<td>0.23</td>
<td>2.6e-2</td>
<td>0.64</td>
<td>3.2e-3</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Figure 4.10 shows the analysis between delay incurred due to the batch size. The delay grows for a growth in batch size. When the value of $k$ is higher, time taken to accumulate the rekeying request is high.
The statistical model discussed in this chapter analyses the behaviour of the user clearly. From the analysis it is clear that the size of the batch rekeying mainly depends on the subscribers activity and delay. Though the optimal rekeying interval is identified by tuning the batch size, the communication complexity can be reduced if the group is structured well. To manage the group, the group can be sub grouped and the next chapter talks about the mathematical model to structure the group.