Future Generation Computer Systems 87 (2018) 259–266

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

RIMS: A Real-time and Intelligent Monitoring System for live-broadcasting platforms

Yangfan Li [a,e], Wei Ren [a,b,e,*], Tianqing Zhu [c], Yi Ren [d], Yue Qin [a], Wei Jie [f]

[a] *School of Computer Science, China University of Geosciences (Wuhan), PR China*
[b] *Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences (Wuhan), Wuhan, PR China*
[c] *School of Information Technology, Deakin University, Australia*
[d] *School of Computing Science, University of East Anglia, UK*
[e] *Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University, PR China*
[f] *School of Computing and Engineering, University of West London, UK*

ARTICLE INFO

Article history:
Received 28 January 2018
Received in revised form 20 March 2018
Accepted 4 April 2018
Available online 17 April 2018

Keywords:
Live streaming platform
Anomaly detection
Fuzzy matching
Frame difference analysis
State awareness

ABSTRACT

Personal live shows on Internet streaming platforms currently are blooming as one of the most popular applications on mobile phones and especially attracting millions of young generation users. The content supervision on live streaming platforms, in which there are thousands or hundreds of show rooms for performing and chatting synchronously, is a major concern with the development of this new service. Traditional image captures and real-time content analysis experience huge difficulties such as processing delay, data overwhelming, and matching overhead. In this paper, we propose a comprehensive method to monitor real-time live stream and to identify illegal or unchartered live misbehaviors intelligently based on various proposed aspects instead of image analysis only. The proposed system called RIMS makes use of several novel indicators on show room status rather than analyzing images solely to support real-time requirements. Three detecting techniques are adopted: self-adaptive threshold-based abnormal traffic detection, sensitive Danmaku comment perception, and frame difference analysis. RIMS can detect dramatically increasing of user number in a show room, filter sensitive words in Danmaku, and capture segmentation of video scenes by frame difference analysis. We deploy our system to monitor a typical live- broadcasting platform called panda.tv, and overall accuracy of detection via three indicators reaches 90.1%. The application of RIMS can change current supervison methods on live platforms that they totally rely on real-time manual review or after the event check. The key techniques in RIMS can also be widely employed in many other mobile applications in edge computing such as video surveillance in Internet of Things and mobile short video sharing.

1. Introduction

The number of live streaming platforms and audiences are both remarkably increasing [1,2] recently. For example, Douyu, the biggest live streaming platform in China, announced that the number of active users per month reached to 1.5 billion. As early as 2015, another live streaming platform, Twitch said it had nearly 100 million visitors and 1.5 million broadcasters per month. Thousands of network anchors are playing games on that platform and real-time communicate with their audiences. A popular form of live video streaming always involves charming ladies. They dance, yak, sing, and engage their audiences via mobile phones or personal computers. Audiences may pay for gifts provided by platforms and donate to performers; Performers can share the profit of gifts in proportion with platforms.

Live broadcast has attracted millions of audiences, and it also imposes great difficulties for real-time management and supervision. Its openness, elapse, and a large amount of online viewers may result in significant security risks [3–5]. For example, an illegal or un-properly broadcasting will result in serious influences or indications, especially when the number of participators in the room is large. In Jun. 2016, a well-known performer plays drag racing, causing a car accident. Thousands of viewers have witnessed the incident. In Oct. 2016, a performer in douyu.tv acts inappropriate nudity. Although the room was recognized by the administrator, it is almost half an hour later.

At present, most live streaming platforms mainly take manual review to identify illegal videos. However, it is not easy to define

[*] Corresponding author at: School of Computer Science, China University of Geosciences (Wuhan), PR China.
E-mail address: weirencs@cug.edu.cn (W. Ren).

https://doi.org/10.1016/j.future.2018.04.012
0167-739X/© 2018 Elsevier B.V. All rights reserved.

specific misbehaviors, and too many patterns of illegal behaviors need to be checked by administrators in a short time simultaneously. To reduce the workload, some companies try to use machine-assisted identification. The main idea is straightforward by analyzing real-time video contents. Some algorithms may have good performance in pornographic video recognition by searching a database consisting of a large number of matching samples. However, this method cannot be applied because the delay is unacceptable in that there are thousands or hundreds of live streaming videos simultaneously, and the detection of pornographic video is like looking for a needle in a bottle of hay. Also, once one pattern is included, many other patterns may occur. The accuracy of video analysis solely is infeasible. Moreover, network performers in live streaming platforms may make use of legal loophole. E.g., only a few seconds of illegal broadcast is released for earning instant profit, the identification system thus may be cheated easily.

To fix above problems, this paper proposes a monitoring system that does not directly analyze real-time video frames but detect abnormal show rooms by indirect factors or indicators. Our system makes use of three tailored design: self-adaptively detecting threshold-based abnormal traffics; monitoring Danmaku (real-time comments) during each live broadcast to discover sensitive perception based on fuzzy matching; splitting suspected video stream into separate scenes and focusing on the room where scene changes. When we comprehensively evaluate above all indicators, the detection efficiency and accuracy will both be improved significantly, and especially, we do not need to capture and process each video streams. The proposed system can also discover new type of illegal patterns because the detection is not solely based on image analysis.

The organization of the paper is as follows: The related work is described in Section 2. Section 3 describes some basic settings for monitoring live video streaming platforms. In Section 4, we describe the proposed system called RIMS and key techniques. In Section 5, we evaluate the system performance through extensive experiments. Finally, concluding remarks and possible future work are mentioned in Section 6.

2. Related work

Preventing illegal video content at live streaming video platform is challenging. The most extensive study falls in image recognition. Felzenschwalb and Huttenlocher [6] proposed a graph-based EGBIS approach, which is a super pixel method. Specifically, images are divided into segmentations which is input into the analysis system. The resulting segments are often called super pixels, which can be used for further analysis to compute certain information about objects in pictures and to recognize the content of videos. However, for real-time video content analysis, it is hard to find a proper algorithm to compute super pixel representations without decreasing the quality of the results. Jochen Steiner and Stefanie Zollmann introduced an incremental super pixels for real-time video analysis [7]. The basic idea of the method is to divide the process of traditional EGBIS segmentation into smaller steps. They improved the segmentation methods that are based on finding minimum cuts in a graph. Some other systems analyze the real-time motion to understand the semantic of video content. Yong Wang et al. [8] proposed a real-time video motion analysis system. They use object detection, object tracking and camera motion understanding to get results. The system balances the computational complexity and analysis performance.

The aforementioned analysis systems target to analyze real-time video. They might have a good performance to some extent. However, live broadcasting platform is more complicated than tradition real-time video. The types of traditional video are relatively fixed, and the publisher usually is a specific organization. While in live broadcasting platforms, there are thousands of live rooms that need to be analyzed at the same time. Everyone can broadcast a live show, as long as she has a personal computer or a mobile phone. Monitoring system may not be able to identify new illegal videos and very likely to miss some of them. For a live streaming platform, one time missing may lead to serious social influences and damage the reputation of platform companies. Moreover, we argue that such image recognition based methods cannot be applied in realtime broadcast platform due to processing delay, and especially when there exist thousands of live streaming videos at the same time.

In general, frame difference method is wildly used in background subtraction [9] and object detection and tracking [10]. This method can accurately extract the main content of an image, and analysis the action of the content [11,12]. This technique can be used for detecting sense changing within a manageable delay.

In live streaming video platform, each room has its own network traffics. Many methods for abnormal traffic detection are proposed [13–16]. Zhengmin Xia et al. proposed a real-time and self-adaptive method for abnormal traffic detection based on self-similarity. It works well for detecting abnormal traffics and unknown attacks.

We will comprehensively adapt above methods in building our experimental system, and focus on the indirected factors such as room status and parameters, instead of relying on image analysis solely. Our monitoring methods present following advantages — much faster detection (or shorter delay), more scalable, more general (in terms of video types), and more accurate.

3. Basic settings

In this section, we briefly describe technical structures of the system. Fig. 1 depicts the Input-Processing-Output (IPO) model of the system.

3.1. Live video streaming cloud and Danmaku

Living broadcast cloud platform is an emerging cloud computing platform in mobile Internet. This cloud needs to support a large number of users for video broadcasting and browsing. As it is convenient for interacting by Danmaku between users and hosts, which flies over the video image during broadcasting, it becomes a popular application in young generation users pervasively once it is released.

Danmaku is a real-time word displaying system that shows audiences' feedback on current live video as multiple lines of moving comments overlaid on the screen. It is always used as a common component in live streaming platforms [17].

3.2. Abnormal traffic detection

Abnormal traffic detection for live streaming videos is challenging. Indeed, the number of live streaming rooms may be huge (e.g., more than ten thousands in douyu.tv) and the detection must be realtime, imposing requirements for lightweight and fast solutions. Although currently there exist a lot of methods to detect abnormal network traffics, most of them is not designed for multimedia traffics as well as not for realtime detection. Tailored design for live broadcast platforms is required, in which self-adaptive threshold abnormal traffic detection is promising. We regard show rooms in platforms as a hub-based network, and traffic detection can be easily accomplished [18,19].

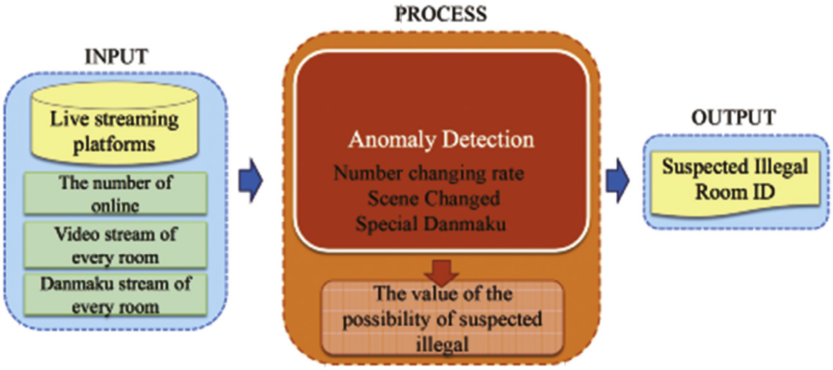
PROCESS
INPUT
Live streaming platforms
The number of online
Video stream of every room
Danmaku stream of every room
Anomaly Detection
Number changing rate
Scene Changed
Special Danmaku
The value of the possibility of suspected illegal
OUTPUT
Suspected Illegal Room ID

Fig. 1. System IPO model.

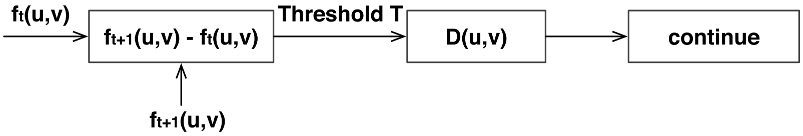

Fig. 2. Frame difference analysis framework.

3.3. *Scenes segmentation*

We use computer image processing algorithms for scene segmentation and variation detection, to finding out whether scenes change in a room. The non-differentiate scene usually presents similar contents, and similar contents can be looked as one detection slot. Live steaming video is composed by many scenes, thus one scene is processed as an element unit. If we can determine that the previous scene is legal and the next scene changes slightly, we can consider that the next scene is also legitimate with high confidence. Consider it in another way, we assume that most illegal scenes are always short in time, are jumped in dramatically for attracting instant influences, and cannot be arranged in advance. Thus, our system focuses on the living rooms which present the obviously scene changing, which can greatly increase the processing speed and reduce the detecting delay.

3.4. *Frame difference analysis*

When any person or object in a live scene moves, a noticeable difference between two adjacent frames could be detected [20]. Frame difference analysis detects adjacent frames of a video to capture the movement. The adjacent two frames will be compared and processed. A live show room is suspected illegal if the computed difference is larger than a tuned threshold. The threshold can be determined and set by an experienced administrator of platforms, and it should be changed dynamically to approach the better false positives and false negatives [21–25].

Frame t subtracts frame $t+1$, and we can get a binary image $D(u,v)$, where u is the row of the pixel, v is the column of the pixel, and the function f is to get the gray value of frame t. $D(u,v)$ is calculated as follows:

$$D(u,v) = \begin{cases} 1, |f_{t+1}(u,v) - f_t(u,v)| \geq T \\ 0, |f_{t+1}(u,v) - f_t(u,v)| < T \end{cases}$$

According to our experience we previously set the threshold as T. In the binary image, 0 represents that there exists no change in the adjacent frames, and 1 represents that there exist changes. The flow chart of the frame difference analysis is shown in Fig. 2.

The advantage of proposed algorithm lies in that various image situations which can be tackled by the adaptively setting parameter.

4. Proposed scheme

Our proposed scheme consists of three primary modules to detect abnormal live show rooms, which constructs several real-time monitor tools for platform administrators. The three modules of our system perform independently, but the ultimate goal is to find suspected illegal rooms. The first module monitors the flow status of a live room per unit time and compares it with our pre-calculated thresholds to judge suspected situations. The second module captures Danmaku flows by connecting Danmaku servers, and then matches Danmaku sensitive words in our pre-set Danmaku library. If the match succeeds, it will prompt suspected violations. The third module relies mainly on the screen-shot from live stream for state sensing and analysis on frame difference, and the suspected illegal room is evaluated by the result of frame difference analysis. In this section, we will introduce above three primary modules. Fig. 3 depicts the architecture of the proposed system.

4.1. *Self-adaptive threshold-based abnormal traffic detection*

The total number of viewers in the room is set as the major indicator of room traffic in this module. In order to implement the abnormal traffic detection by self-adaptive threshold, we conduct technical steps as follows:

Firstly, we use *jsoup*, a HTML parser based on Java, to parse out the number of live rooms and the number of entire online users accordingly from acquired html.

Subsequently, we record the total number of viewers (denoted as nov) in a live room i (i is the room id) for each time segment (denoted as ts).

After that, the growth number per hour, denoted as K, is calculated by using linear regression equation. The formula for K is:

$$K = \frac{\sum_{i=1}^{n} ts_i * nov_i - n\overline{ts * nov}}{\sum_{i=1}^{n} ts_i^2 - n\overline{ts}^2}$$

where n represents the total number of time segments.

Therefore, we define K/N (N is the number of online users at the beginning of this time slot, i.e., hour), which represents the growth rate of online user numbers in this hour.

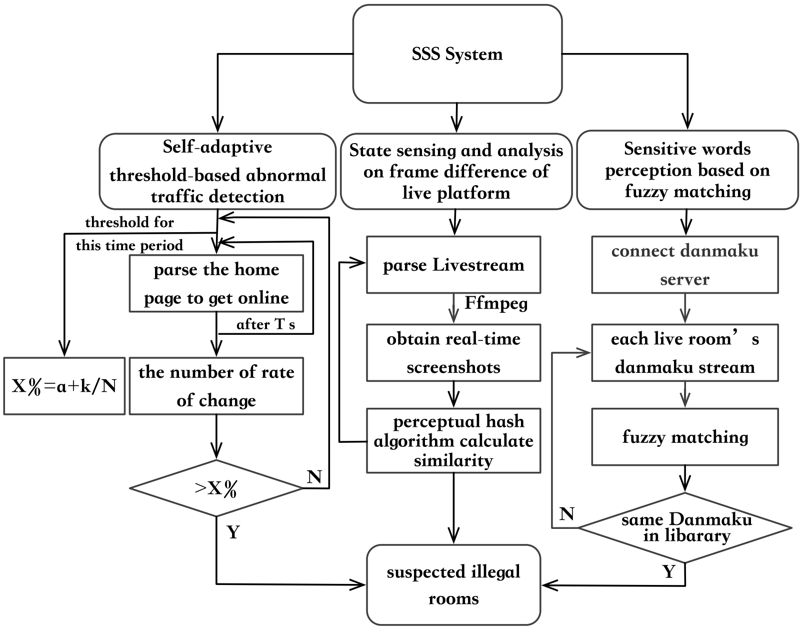
SSS System
Self-adaptive threshold-based abnormal traffic detection
State sensing and analysis on frame difference of live platform
Sensitive words perception based on fuzzy matching
threshold for this time period
parse the home page to get online
after T s
X%=a+k/N
the number of rate of change
>X%
N
Y
parse Livestream
Ffmpeg
obtain real-time screenshots
perceptual hash algorithm calculate similarity
connect danmaku server
each live room's danmaku stream
fuzzy matching
same Danmaku in libarary
N
Y
suspected illegal rooms

Fig. 3. The architecture of the proposed system.

We define $B = 1/N$, and the number of growth rate $K * B$ as an important parameter of adaptive threshold. Thus, we use $K*B+\Delta$ to calculate the threshold value (our observation and experimental experiences convince the rationality of this threshold). The threshold is therefore calculated as follows: $X\% = K * B + \Delta$

If the actual growth rate of online user number is greater than the threshold, this room is regarded as an abnormal room with high confidence, and the number of this room will be displayed on the alter interface of our monitoring system.

We noticed that $K * B$ becomes larger when the number of the online users grows extremely fast during a given time slot. The threshold value can be tuned larger correspondingly. Thus, the actual growth rate of online user number calculated by the system should be bigger for further detection.

On the contrary, if the number of online users grows slowly, or even declines, the threshold value should be set smaller accordingly. Subsequently, when the actual growth rate calculated by system is slightly greater, this room will be considered as abnormal. Hence, this adaptive threshold matchs with the most recent status.

In other words, the system tunes a threshold value related to the current time and room id, which is a normal rate and denoted as X, for each live room about online number per hour. The threshold is adaptive as realistic situations of each time may be differentiate, in order to avoid false positive and false negative jitters. Therefore, the threshold value should fall in an acceptable range of deviation. The adaptive algorithm for tuning of threshold value is given in Algorithm 1 as follows:

Algorithm 1: Self-adaptive threshold-based abnormal traffic detection

```
Input: current time , living platform's target url
Output: unusually $roomid_i$
Data:
$n \leftarrow$ living platform's room number
$K \leftarrow$ the rate of change of all people
$X \leftarrow threshold$
for $i \leftarrow 1$ to $n$ do
    $xi \leftarrow$ the current number of people in the $Room_i$
    $yi \leftarrow$ the number of people in the $Room_i$ after t seconds
    $K = \frac{\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$
    $X\% = a\% + KB$
    if $K > X$ then
        output unusual $roomid_i$
    else
        continue
    end
end
```

4.2. Sensitive word perception

This module is achieved by simulating multiple clients that can connect Danmaku servers to elicit Danmaku streams. As for the detection on sensitive messages in flying Danmaku, the traditional way only concentrates on sensitive words from huge amounts of Danmaku databases, which reduces detection efficiency. Thus, we first collect large amounts of sensitive words in the scope of anti-terrorism, heresy, pyramid selling, disunion and pornography, etc. Subsequently, we list some possible keywords based on the frequency of sensitive words for this room. To reduce the computation overhead, only fuzzy matching is conducted in analyzing of Danmaku. The system selects key words from Danmaku streams, and then matchs them with the words in a keyword table. If it matches, the room number and the information of the Danmaku sender will be displayed on the alter screen of the monitoring system.

We use a customized KMP algorithm to implement the sensitive Danmaku message fuzzy matching. KMP is a string-matching algorithm with high efficiency and agile implementation. Additionally, this algorithm costs the shortest time, compared with algorithms of the same type. KMP algorithm makes full use of the information contained in a specific pattern, to obtain a prefix module by preprocessing this pattern. This algorithm performs more efficiently than traditional ones.

KMP algorithm to fuzzy match the information of Danmaku can find out similar words and phrases related to keywords (for example, the keyword is "shit". If there exist "sh.. it*, s... hit, shi... tttttt" in the Danmaku stream that are similar to the key words, they will be identified.) Besides, it can also find out some representatives or typical keywords which are used by message senders who are incentive to avoid being detected. All of them can highly improve the efficiency and accuracy of detection. Also, the match method transforms Chinese characters to pinyin. Consequently, they can match with each other without being the same intonation. This match improves the traditional way, such as char-pattern matching and calling library, by exalting both matching speed and accuracy. The proposed method is given in Algorithms 2 and 3 as follows:

Algorithm 2: Sensitive words perception based on fuzzy matching

```
Input: living platform's target url
Output: unusually $roomid_i$
Data:
$X \leftarrow threshold$
$n \leftarrow$ living platform's room number
$K \leftarrow$ the rate of change of all people

build danmu library Including counter-terrorism,cults,etc.
for $i \leftarrow 1$ to $n$ do
    danmu stream $\xleftarrow{get}$ link danmu server
    if danmu string s1 is not NULL then
        string t1 $\xleftarrow{pinyin}$ string s1
        if fuzzy-matching(t1) is TRUE then
            output unusual $roomid_i$
        else
            continue
        end
    else
        continue
    end
end
```

Algorithm 3: fuzzy-matching(t1) module

```
Input: text, pattern
Output: 0 or 1

$j \leftarrow 0$
$k \leftarrow 0$
for $i \leftarrow 1$ to $n$ do
    while j and p[j]!=t[i] do
        j=f[j]
    end
    if $p[j] == t[i]$ then
        j++
    end
    if $j == m$ then
        k++
    end
end
return k
```

4.3. State sensing and analysis on frame difference

This method is achieved through the following steps. First, we capture live video stream addresses from live URL, and then set a time interval T. After that, we call *ffmpeg* command to obtain screen-shots of the video within the same time interval. Subsequently, we name the screen-shots according to a special naming rules and save it to the local disk. We tailored design perception of hash perceptual [26,27] algorithm to calculate similarity between consecutive screen-shots. When the live room id is altered and the similarity gaps is sufficiently large, the monitor system guides attentions to the live room with lower similarity. The low similarity implies that room state changes remarkably, and the supervisor will be notified to check it.

The scene segmentation technique is depicted in Fig. 4.

The proposed system detected that scene changed in room 1 at time *t1*, and no changes at room 2 and room 3. At time *t2*, the scenes of room 1 and room 2 changed; that of room 3 did not change. Room 1 at *t1* thus needed more attentions, likewise were room 1 and room 2 at *t2*. We divided the live stream scenes via adjacent frame differences. Image perceptual hashing algorithm is implemented for state perception analysis onto frame difference, which relies on information processing theory from cognitive psychology. A mapping is created from multi-media data set to multi-media perception set, which satisfies the demand for perception security. The revised algorithm calculates the similarity between consecutive image that feeds to perception hash algorithm, which has multiple advances such as robustness, discrimination, collision resistance and unidirectionality. The details are presented in Algorithm 4.

Algorithm 4: State sensing and analysis on frame difference of live platform

```
Input: living platform's target url
Output: unusually $roomid_i$
Data:
$Y \leftarrow$ similarity threshold
$n \leftarrow$ living platform's room number

for $i \leftarrow 1$ to $n$ do
    build ffmpeg command
    save 1.bat
    run 1.bat
    get roomid_t.bmp
    if roomid_t-1.bmp is exit then
        $y \leftarrow$ calculate the similarity
        if $y < Y$ then
            output unusual $roomid_i$
        else
            continue
        end
    else
        continue
    end
end
```

We compute two-dimensional Discrete Cosine Transformation (DCT) in perceptual hashing algorithm [24,28] to explore frequencies and amplitudes for approximating an image. Two-dimensional DCT is divided into positive and inverse transform, and major operations are listed as follows:

1. Positive transformation:

$$F(u, v) = \alpha(u)\alpha(v)\sum_{x=0}^{P-1}\sum_{y=0}^{Q-1} f(x, y) cos\frac{u\pi x + \frac{1}{2}u\pi}{P} cos\frac{v\pi y + \frac{1}{2}v\pi}{Q}$$

$u \in [0, P-1],\ v \in [0, Q-1]$

$$\alpha(u) = \begin{cases} \sqrt{\dfrac{1}{P}}, & u = 0 \\ \sqrt{\dfrac{2}{P}}, & u \neq 0 \end{cases} \qquad \alpha(v) = \begin{cases} \sqrt{\dfrac{1}{Q}}, & v = 0 \\ \sqrt{\dfrac{2}{Q}}, & v \neq 0 \end{cases}$$

2. Inverse transformation:

$$f(x,y)=\sum_{x=0}^{P-1}\sum_{y=0}^{Q-1}\alpha(u)\alpha(v)F(u,v)cos\frac{u\pi x+\frac{1}{2}u\pi}{P}cos\frac{v\pi y+\frac{1}{2}v\pi}{Q}$$

$$u\in[0,P-1],\ v\in[0,Q-1]$$

$$\alpha(u)=\begin{cases}\sqrt{\frac{1}{P}}, & u=0\\ \sqrt{\frac{2}{P}}, & u\neq 0\end{cases}\qquad \alpha(v)=\begin{cases}\sqrt{\frac{1}{Q}}, & v=0\\ \sqrt{\frac{2}{Q}}, & v\neq 0\end{cases}$$

5. Experiment and performance evaluation

The system relies on the harmonic cooperation of three components — abnormal traffic detection, sensitive word perception based on fuzzy matching, and perception state analysis on frame difference. Accordingly, in order to show the efficiency of the system, we first test these three parts separately, and then conduct an overall study. We evaluate multiple performances in real experiments. From Sept. 20th to 27th, we have tested more than 100 rooms in Panda.tv for 7 days. We chose this live platform because it is one of the biggest live streaming platform in China, and we can easily get the detail information of the site. All the following experiments are conducted by Java over PC with Intel Core i5 wit 2.50 GHz processor and 4GB memory.

(1) Abnormal traffic detection.

It finds the suspected illegal rooms by the rate of traffic changes, which is reflected by the changing rate of audience number in the room. Our experiment evaluates two folders: all rooms with abnormal traffic could be detected; how many detected rooms are really suspected illegal. The details are given in the following.

First, we test whether all rooms with abnormal traffics could be detected by the system. For each 10 minutes, our system fetches the number of audiences in each room, and calculates the changing rate of number in adjacent intervals for each room. The system records the results in a log file, including threshold, links of each room, on-line numbers of each room, and changing rate of viewer number in adjacent intervals for each room. The results are shown in Fig. 5. It is easy to check whether rooms with abnormal traffics can be detected by the system. As shown in the figure, our system will focus on room 352783 at 10 min and room 13653 at 100 min.

Secondly, we check the rooms that are detected by the system to verify whether they are really suspected illegal. It can be done by opening windows of these rooms manually. During our experiments, there are 69 rooms with abnormal traffic detected by the system, and 32 of them are really suspected illegal. Observing from experiments, we find that if a room contains sensitive words, such as relating to pornography, the number of audiences will sharply increase, which influences on traffics. This again justifies that abnormal traffics can serve as a reference for the detection of suspected illegal rooms.

(2) Sensitive words perception.

This part detects suspected illegal rooms via Danmaku. The sensitive words in Danmaku streams include following categories: terrorism, racial discrimination, religion, pornography, dirty words, and so on. If Danmaku contains sensitive words, the room might be suspected illegal. A library of sensitive words is constructed. In our experiments, Danmaku of each room is monitored. If Danmaku in a room contains sensitive words, related records will be dumped in a log file. We then check that in these rooms how many are really suspected illegal. In detection, 102 rooms are found sensitive words, of which 45 are really suspected illegal rooms. In the experiments, we found that Danmaku is an indicator of interaction hotness between audiences and a host. The current number of flying Danmaku is also related to show contents and the number of audiences. Therefore, sensitive words can be used as a reference for suspected illegal rooms detection.

(3) State of perception and analysis on live frame difference.

Frame difference can be applied to scene segmentation. By capturing moving objects in the same intervals and analyzing the image sequence of adjacent frames, we can detect those rooms where the frame differences change rapidly so as to be larger than

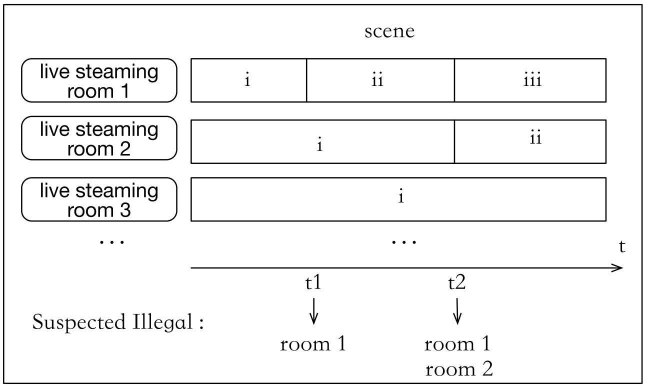

Fig. 4. Traffic monitoring framework.

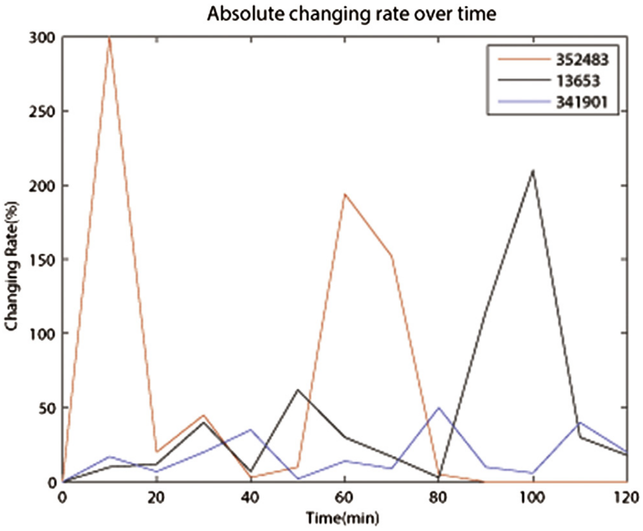

Fig. 5. Log file for traffic monitoring results.

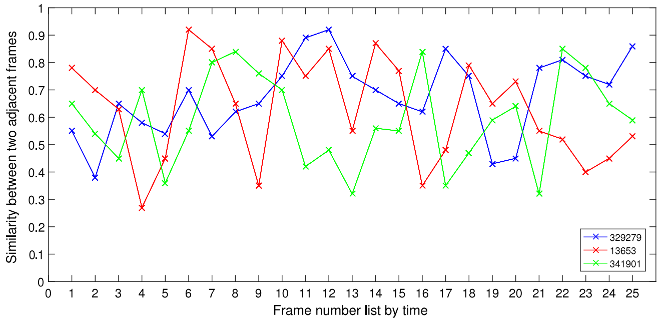

Fig. 6. The changes of frame difference of Room 329279.

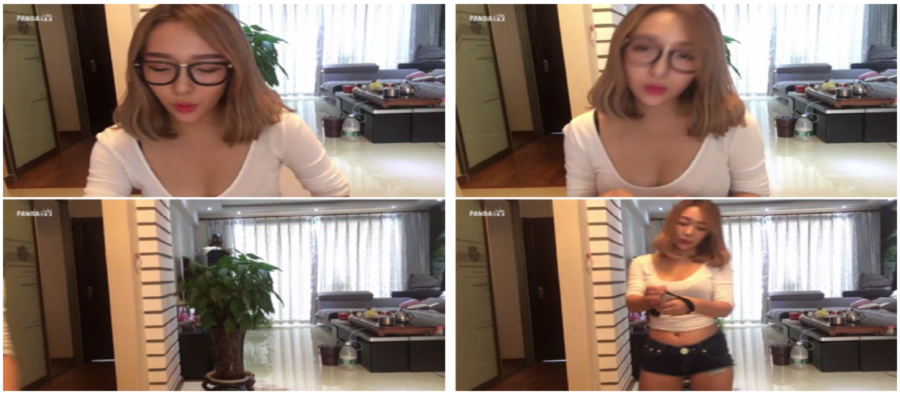

Fig. 7. The screen-shots of Room 329279 - 18 frames to 21 frames.

Table 1
Detection results by RIMS (live streaming rooms in Panda.tv).

Detection indicator	Output room number	Suspected illegal rooms	Accuracy
Abnormal traffic detection	69	22	31.8%
Sensitive words perception	102	45	44.1%
State of perception and analysis	257	76	29.6%
Union three indicator	43	39	90.1%

the threshold. These rooms are very likely suspected illegal rooms, thus we could use it as an indicator.

To test this module, we capture screen-shots of each room every 10 seconds and calculate the frame difference of adjacent intervals for each room. If the frame difference is above the threshold, our system will alter the room id. We select room 329279 in panda.tv as an example. This room is detected by abnormal traffics, and it is often recommended on the web site. The changes of frame difference are shown in Fig. 6. Fig. 7 presents the adjacent screen-shots of room 329279. The experiment result again justifies that state perception and analysis on frame difference can serve as an indicator for suspected illegal room detection.

(4) Comprehensive analysis

Based on aforementioned three detection methods, we detect live streaming rooms in Panda.tv. We select those suspected illegal rooms from output rooms by manual detection, then record data in Table 1.

Extensive experiment analysis shows that above three proposed methods are important indirect indicators for detecting suspected illegal rooms, and our system comprehensively takes advantages of them all for optimized efficiency and accuracy.

6. Conclusion and future works

In this paper, we proposed a real-time intelligent monitoring system called RIMS for live video streaming platforms based on three novel indirect indicators such as state awareness, Danmaku filtering, and frame difference analysis. The proposed scheme comprehensively considers abnormal traffic in terms of changing rate of room audiences, sensitive words filtering in Danmaku, and state perception by frame difference analysis. We evaluate our system in a well-known platform Panda.tv. We also recorded the number of suspected illegal live streaming, which was selected by manual detection (online review manually). As a result, our system alerted 43 rooms that detected by our proposed factors, and 39 of them are really suspected illegal. The integrated accuracy of three modules in RIMS can reach 90.1%. The experiments show that our proposed methods significantly reduce working load of content supervision on current live streaming platforms as well as guarantee real-time alter and repsonse, in which most of them rely on manually monitoring or after the event checking.

Acknowledgments

The research was financially supported by the Major Scientific and Technological Special Project of Guizhou Province (20183001), Open Funding of Guizhou Provincial Key Laboratory of Public Big Data (2017BDKFJJ006), National Science Foundation China (61502362), and Open Funding of Hubei Provincial Key Laboratory of Intelligent Geo-Information Processing (KLIGIP2016A05).

References

[1] J. Peng, S. Detchon, K.-K.R. Choo, H. Ashman, Astroturfing detection in social media: a binary n-gram-based approach, Concurr. Comput.: Pract. Exper. 29 (17) (2016) e4013.

[2] A. Heravi, D. Mani, K.-K.R. Choo, S. Mubarak, Making decisions about self-disclosure in online social networks, in: Proceedings of 50th Annual Hawaii International Conference on System Sciences, (2017) 1922–1931.

[3] D. Quick, K.-K.R. Choo, Pervasive social networking forensics: Intelligence and evidence from mobile device extracts, Netw. Comput. Appl. 86 (2017) 24–33.

[4] C.-D. Orazio, K.-K.R. Choo, An adversary model to evaluate DRM protection of video contents on iOS devices, Comput. Security. 56 (2016) 94–110.

[5] S.A. Miraftabzadeh, P. Rad, K.-K.R. Choo, M.A. Jamshidi, Privacy-aware architecture at the edge for autonomous real-time identity re-identification in crowds, IEEE Internet Things J. (2017) 1–1.

[6] P.F. Felzenszwalb, D.P. Huttenlocher, Efficient graph-based image segmentation, Int. J. Comput. Vis. 47 (2004) 167–181.

[7] J. Steiner, S. Zollmann, G. Reitmayr, Incremental superpixels for real-time video analysis, in: Proceedings of the Computer Vision Winter Workshop. 164 (2011) 368.

[8] Y. Wang, T. Zhang, D. Tretter, Real time motion analysis toward semantic understanding of video content, in: Conference on Visual Communications and Image Processing. 5960 (2005) 596027.

[9] P. Ramyaa, R. Rajeswarib, A modified frame difference method using correlation coefficient for background subtraction, Proc. Comput. Sci. 93 (2016) 478–485.

[10] S. Kamate, N. Yilmazer, Application of object detection and tracking techniques for unmanned aerial vehicles, Proc. Comput. Sci. 61 (2015) 436–441.

[11] G.-S. Sarma, V.-S. Kumar, S. Nizmi, Image processing - A gizmo for video content extraction, ARPN J. Syst. Softw. 1 (8) (2011).

[12] M. Jiang, J. Kong, H. Huo, Informative joints based human action recognition using skeleton contexts, Signal Process. Image Commun. 33 (2015) 29–40.

[13] W. Xiong, H. Hua, N. Xiong, L.-T. Yang, W.-C. Peng, X.-F. Wang, Anomaly secure detection methods by analyzing dynamic characteristics of the network traffic in cloud communications, Inf. Sci. 258 (2014) 403–415.

[14] V.-S.W. Eide, O.C. Granmo, F. Eliassen, Real-time video content analysis: QoS-aware application composition and parallel processing, ACM Trans. Multimedia Comput. Commun. Appl. 2 (2) (2006) 149–172.

[15] M.S. Kim, H.J. Kong, S.C. Hong, A flow-based method for abnormal network traffic detection, Netw. Oper. Manag. Symp. 1 (2004) 599–612.

[16] Z. Xia, S. Lu, J. Li, Real-time and self-adaptive method for abnormal traffic detection based on self-similarity, in: International Conference on Web Information Systems and Mining, Springer Berlin Heidelberg, (2009) 383–392.

[17] Y. Chen, Q. Gao, P.-L.P. Rau, Watching a movie alone yet together: Understanding reasons for watching danmaku videos, Int. J. Human-Computer Interact. 33 (9) (2017) 731–743.

[18] W. Ren, S. Huang, Y. Ren, K.-K.R. Choo, Lipisc: a lightweight and flexible method for privacy-aware intersection set computation, Plos One 11 (6) (2016) e0157752.

[19] W. Ren, R. Liu, M. Lei, K.-K.R. Choo, Segoac: a tree-based model for self-defined, proxy-enabled and group-oriented access control in mobile cloud computing, Comput. Stand Interfaces 54 (2016) 29–35.

[20] X.-H. Jin, X. Hui, W. Yang, Q.-J. Liu, D. Zhao, S. Xu, Improved moving target detection technology, Adv. Mater. Res. 971 (2014) 1628–1632.

[21] H. Liu, H. Kun, J.-L. Dai, R. Wang, H.-Y. Zheng, B. Zheng, Combining background subtraction and three-frame difference to detect moving object from underwater video, OCEANS (2016) 1–5.

[22] Y.-H. Miao, L.-Z. Wang, D.-S. Liu, Y. Ma, W.-F. Zhang, L.-J. Chen, A. Web, 2.0-based science gateway for massive remote sensing image processing, Concurrency Comput. Practice Exp. 27 (9) (2015) 2489–2501.

[23] L.-Z. Wang, Y. Ma, J.-N. Yan, V. Chang, A.-Y. Zomaya, pipsCloud: High performance cloud computing for remote sensing big data management and processing, Future Gen. Comput. Syst. 78 (2018) 353–368.

[24] L.-Z. Wang, W.-J. Song, P. Liu, Link the remote sensing big data to the image features via wavelet transformation, Cluster Comput. 19 (2) (2016) 793–810.

[25] Y. Liu, Y. Zhao, M. Liu, L. Dong, Y. Wu, Research on the algorithm of infrared target detection based on the frame difference and background subtraction method, Signal Data Process. Small Targets 9596 (2015) 959607.

[26] N. Saikia, P.-K. Bora, Perceptual hash function for scalable video. International journal of information security, J. Real-Time Image Process. 13 (1) (2014) 81–93.

[27] L. Weng, B. Preneel, A secure perceptual hash algorithm for image content authentication, in: IFIP International Conference on Communications and Multimedia Security, Springer Berlin Heidelberg, 2011 (108–121).

[28] Q.-W. Zhang, W.-X. Du, L.-Q. Yuan, M. Li, Face recognition using discrete cosine transform and fuzzy linear discriminant analysis, Commun. Comput. Inf. Sci. 244 (2011) 286–293.

Yangfan Li is a master student at School of Computer Science, China University of Geosciences (Wuhan), China. His research interests include information security and image processing.

Wei Ren currently is a Professor in School of Computer Science, China University of Geosciences (Wuhan), China. He was with Illinois Institute of Technology, USA in 2007 and 2008, School of Computer Science, University of Nevada Las Vegas, USA in 2006 and 2007, and Hong Kong University of Science and Technology, in 2004 and 2005. He obtained his Ph.D. degree in Computer Science from Huazhong University of Science and Technology, China. He published more than 70 refereed papers, 1 monograph, and 4 textbooks. He obtained 10 patents and 5 innovation awards. He is a senior member of China Computer Federation.

Tianqing Zhu received her B.Eng. and M.Eng. degrees from Wuhan University, China, in 2000 and 2004, respectively, and a Ph.D. degree from Deakin University in Computer Science, Australia, in 2014. Dr. Tianqing Zhu is currently a continuing Teaching Scholar in the School of Information Technology, Deakin University, Australia. Her research interests include privacy preserving, data mining and network security. She has won the best student paper award in PAKDD 2014.

Yi Ren obtained his Ph.D. in Information Communication and Technology from the University of Agder, Norway in 2012. He was with the Department of Computer Science, National Chiao Tung University (NCTU), Hsinchu, Taiwan, as a Postdoctoral Fellow and an Assistant Research Fellow from 2012 to 2017. He is currently a Lecturer in the School of Computing Science at University of East Anglia (UEA), Norwich, U.K. His current research interests include security and performance analysis in wireless sensor networks, ad hoc, and mesh networks, LTE, and e-health security. He received the Best Paper Award in IEEE MDM 2012.

Yue Qin is a master student at School of Computer Science, China University of Geosciences (Wuhan), China. Her research interests include mobile security and authentication of mobile devices.

Wei Jie has been active in a broad spectrum of areas in parallel and distributed computing, in particular, grid and cloud computing, computing security technologies, e-science and e-research. Dr. Jie has been actively involved in professional services. He is the General Chair of the IEEE workshop on Security in e-Science and e-Research, and has served as Program Committee member for more than 40 international conferences and workshops. He has published approximately 50 papers in international journal and conferences and has edited three books.