Suspicious Loitering Detection from Annotated CCTV Feed Using CEP Based Approach

Pengesanan Pergerakan Mencurigakan dari Suapan CCTV Teranotasi Menggunakan Pendekatan CEP

ABSTRACT

Smart Surveillance System is a critical system that enables automated detection of anomalous activities from live CCTV feed. The main challenge that needs to be addressed by the Smart Surveillance System is the ability to understand and detect the activities that are currently occurring within the CCTV feed. Suspicious loitering is considered one of the anomalous activities that precede unwanted events, such as break-ins, burglary, and robbery. In this research, the Complex Event Processing (CEP) approach was selected as the system development approach for developing a Smart Surveillance System. Four types of similarity search-based event detectors, namely the Multi-Layered Event Detector for General Application (MEGA), Temporally Constrained Template Match Detector (TCD), Sliding Window Detector (SWD), and Weighted Sliding Window Detector (WSWD) were tested and evaluated to determine the best suspicious loitering event detector to be used in the Smart Surveillance System. The input data to the detectors comprised manually annotated real CCTV feed which was subjected to three noise conditions: (i) no-noise (0% noise) annotation, (ii) 25% noisy annotation and (iii) 46.8% noisy annotation. The 46.8% noisy annotation is assumed to reflect the real ambient operating condition of the Smart Surveillance System; while the no-noise condition was assumed to reflect the perfect CCTV feed acquisition and annotation process. The performance of the detectors was measured in terms of sensitivity, specificity, detection accuracy, and the area under the Receiver’s Operating Curve (ROC). The results obtained showed that MEGA is the best overall detector for suspicious loitering detection in ambient operating conditions with detection accuracy of 97.20% and area under ROC curve of 0.6117.

Keywords: Event detection; smart surveillance system; complex event processing

ABSTRAK

Sistem Pengawasan Pintar adalah sebuah sistem kritikal yang membolehkan pengesanan aktivitianomali dilakukan daripada suapan CCTV dalam talian. Cabaran utama yang perlu diatasi oleh Sistem pengawasan Pintar adalah keupayaannya untuk memahami dan mengesankan aktiviti-aktiviti yang sedang berlaku di dalam suapan CCTV. Pengesanan mencurigakan adalah salah satu daripada aktiviti anomali yang mendahului peristiwa-peristiwa yang tidak diingini seperti pecah masuk, kecurian dan rompakan. Di dalam kajian ini, pendekatan Pemprosesan Peristiwa Kompleks (CEP) telah dipilih sebagai pendekatan pembangunan sistem bagi membangunkan sebuah Sistem pengawasan Pintar. Empat jenis pengesan peristiwa berasaskan pencarian keserupaan, iaitu, Pengesan Pemadanan terkekang Temporal (TCD), Pengesan Tetingkap Menggelongsor (SWD), Pengesan Pemadanan terkekang Temporal (TCD), Pengesan Tetingkap Menggelongsor (SWD) dan Pengesan Tetingkap Menggelongsor dengan sub-Tetingkap Berpemberat (WSWD) telah diuji dan dilimpah bagi menentukan pengesanan peristiwa pergerakan mencurigakan terbaik untuk digunakan di dalam Sistem Pengawasan Pintar tersebut. Data masukan kepada pengesan-pengesan tersebut adalah suapan CCTV sebenar yang telah dianotasi secara insani dan telah dikenakan dengan tiga keadaan hingar : (i) anotasi tanpa hingar (0% hingar), (ii) anotasi dengan hingar sebanyak 25% dan (iii) anotasi dengan hingar sebanyak 46.8%. Anotasi dengan hingar sebanyak 46.8% dianggap mewakili keadaan persekitaran sebenar bagi Sistem Pengawasan Pintar manakala keadaan tanpa hingar pula dianggap mewakili proses pemerolesan suapan CCTV dan anotasi yang sempurna. Prestasi pengesan diukur di dalam bentuk sensitiviti, spesifisiti, ketepatan pengesanan dan kawasan dibawah Keluk Operasi Pengguna (ROC). Keputusan yang diperolehi menunjukkan bahawa MEGA adalah pengesanan keseluruhan terbaik bagi pengesanan pergerakan mencurigakan di dalam keadaan persekitaran sebenar dengan ketepatan pengesanan sebanyak 97.20% dan kawasan dibawah keluk ROC sebanyak 0.6117.

Kata kunci: Pengesan peristiwa; Sistem pengawasan pintar; Pemproses peristiwa kompleks
INTRODUCTION

The smart surveillance systems are growing in demand as crimes are quite rampant nowadays. Current smart surveillance system uses CCTV as means of monitoring, investigating and preventing crimes. Generally, the smart surveillance system is used to monitor, identify and describe patterns of events that indicate odd behavior of human activity or to be specific, anomalous behavior. Although the present smart surveillance system is equipped with functions to collect, store and send information, the existing smart surveillance system still lacks the capability to recognize and identify events that are currently occurring within the CCTV (Hilal et al. 2011). Studies have shown that the performance of a security guard will be significantly reduced after focusing on the screen which displayed video CCTV for 20 minutes (Castro et al., 2011). In addition, the operator or security personnel also failed to correlate every event through video footage obtained from various CCTVs in a building to come out with a unified understanding of what event is currently happening in the monitored environment. As such, this situation indirectly contributes to an increasing number of potential mishaps. Smart surveillance or CCTV systems also have limitations as they have to be operated manually and are not able to detect and recognize anomalies such as robberies and intrusions automatically. Furthermore, a centralized CCTV monitoring system requires an operator to be in a controller room or area that is far from the areas being monitored (Shahad et al. 2016). When anomalies are detected from the CCTV feed, the operator must leave the controller room to carry out check on the spot, thus decreasing the surveillance system effectiveness. These limitations reduce the efficiency of a CCTV based surveillance system.

Recently, smart surveillance systems have started to get the attention of the public to combat crimes. Increasing number of crimes had brought awareness to the public that the need for the smart surveillance system at their premises is a must to ensure the personal safety and property. Features such as high-speed data transmission, wireless networks and remote surveillance cameras can be applied in the development of such smart surveillance system based on real-time events to monitor a variety of conditions and events that took place in a building (Rakesh et al. 2012). Network-based control system has been adopted in the intelligent building system in recent years. Such system enables network-based real-time monitoring of a building and also facilitates the management of a building automation system that systematically collect, analyze and store the information on the building. In this context, the above technology can be applied in the smart surveillance system to facilitate the recognition and object tracking processes. In the past (Adi et al. 2006), Complex Event Processing (CEP) technology has been applied in the business world, especially in the banking sector and the insurance industry to monitor and detect certain conditions, such as suspicious money transfers, fraud detection and market trends. This paper proposes a loitering event detection application based on annotated CCTV feed. This approach allows the creation of unauthorized loitering or suspicious movement detection in monitored areas. In this work, CEP will be applied to detect the occurrence of an event by evaluating the observed sequence of events based on the rules set out in the CAISER (Complex Event Processor for Scientific and Engineering Application) platform.

LITERATURE REVIEW

Smart surveillance systems are getting increasingly popular. There are many studies have been made towards the application of this system. For example, the development of home surveillance application and smart surveillance for generic purpose (Ali 2016; Antoniou & Angelov 2016), monitor displaced activity detection system (Babutain et al. 2015) and smart surveillance system for secure environment from fight, riots of violence protests, panic behavior and excitement (Fookes et al. 2010). However, the most challenging part in developing smart surveillance system is to develop the event detection capability so that the smart surveillance system can identify what is happening and take mitigation action accordingly. There are several algorithm and approach to execute the above automatically. One of the most promising approach to do that is through the use of CEP approach.

In CEP system, Simple Events (SE) are acquired from event generators, processed centrally and ComplexEvents (CE) which are deduced from the interaction of current and previous SE with previous CE under pre-determined temporal and spatial constraint. The system then executes the mitigation action automatically for selected events. CEP has emerged as a new paradigm in solving this problem by taking the role as a middleware, providing significant reports and improving the system automation through filtering, aggregating and constructing the complex events (Zacheilas et al. 2015). CEP is an emerging research field that arouses researches attention in the last decade. Li & Bingwen(2010) introduced a CEP system that deploys network nodes computing power to reduce traffic waste due to communication data for a wireless sensor network (WSN). The event processing queries, server and nodes' queries allow local real-time event processing which help to save network traffic and reduce collision chance. Wang et al. (2013) proposed a high performance hierarchical probabilistic CEP (pCEP) methodology that utilizes both extended Nondeterministic Finite Automation (NFA) model and a tree-based query model to support complex event processing over single distributed event stream. This data mining approaches work effectively over distributed event stream with large sliding window size, but CEP is only constricted for NFA based pCEP engine due to immature optimization of query model. Saleh and Sattler (2013) suggested an application of In-Network Distributed CEP (INDECP) to perform complex event processing in a sensor network before disseminating them to destinations. An early stage in-network NFA-based complex event detection
system with low data transferring energy was developed by them for acquiring meaningful information from the sensor network.

A brief survey of relevant literature shows that most researchers emphasize on the CEP design query, event processing language and optimization of real time analysis whereas the systems normally utilize the experts devised rule patterns (Mehdiyev et al. 2015). The lack of Machine Learning (ML) in present research works can be improved by incorporating CEP middleware with data mining approaches (Mehdiyev et al. 2015). Several ML detectors have been implemented in domains such as intrusion detection and fingerprint identification for interest event patterns identification from sensor data streams (Lee et al. 1999). Shah et al. (2010) implemented sliding window (SW) method with varying window size on wireless sensor nodes for scalar multiplication. This method can considerably reduce the node failure risk and at the same time promotes memory-saving. Kavitha & Suresh (2012) presented an effective intrusion prevention system with time-sensitive data mining approach which is aimed to assist in real time high speed massive data stream processing. Results showed that the time-sensitive SW performs better than correlation property in processing enormous data stream. However, the proposed approach considers merely for a simple attack event due to slow computational processing. Previous work on CEP based smart surveillance was also done by Saad (2017). The results shown that the approach of CEP is very promising for smart surveillance system. For a smart surveillance system, CE detection basically comprises several activities which are video events detection, target features description as well as semantic concept finding (Ke et al. 2016). An example of commercially available Smart Surveillance System is shown as in Figure 1.

![Smart Surveillance System](source)

**FIGURE 1.** The architecture of a commercially available Smart Surveillance System

**Source:** www.puretechsystems.com

**METHODOLOGY**

The general objective of this research is to develop a loitering event detection application using the CEP approach. The experiments for this research were conducted around the Innovation Laboratory 2, Faculty of Engineering & Built Environment, UKM. There are 3 major steps involved in the experiment which are (1) Data Collection using CA-InViSSLite, (2) Video Annotation using InViSS Annotator and finally (3) Complex Event detection using CA-CED, CA-InViSSLite, InViSS Annotator and CA-CED are CCTV and surveillance based CEP tools developed previously by Saad (2017). The process is summarized in Figure 2.
FIGURE 2(a). Video data acquisition via CA-InViSS.Lite, (b) Video event annotation using InViSS Annotator and (c) Complex Event Detection using CA-CED
There are four regions of interest (ROIs) that have been selected for event detection testing purposes for this experimental study as shown in Figure 3. Figure 3(a) depicts the schematic layout around the testing areas including main entrance area (green pattern) and corridor area (yellow pattern). Figure 3(b) depicts the actual layout that includes the ROI of R1, showing a parking area around main entrance, ROI of R2, showing the entrance outside building and ROI of R3, and R4 showing the corridor area inside the building.

A CCTV based surveillance system was used for real time monitoring of the above ROI. Figure 2(a) depicted the monitored environment videos captured using CA-InViSSLite. The detected events are displayed on the event log whenever there are movement on the designated ROI.

**VIDEO ANNOTATION**

The recorded videos are then transferred into InViSS Annotator which is a video annotation tool that can be used to annotate SE and CE for each frame in the recorded video. InViSS Annotator is used to generate events sequence file automatically based on the annotations. Figure 2(b) depicts the video annotation process. In this research, a total of 500 annotations were made for the test video sequence. Table 1 shows the types of noise introduced. The annotations were stored in 3 sequence files with varying level of artificially induced noise: file 1 with 0% noise representing the perfect data acquisition condition, file 2 with 25% noise representing a low noise condition and file 3 with 46.8% noise representing a data acquisition processed polluted with noise typical for a real ambient operating condition.

**TABLE 1. Types of noise introduced**

<table>
<thead>
<tr>
<th>Type of Noise</th>
<th>Annotated Events</th>
<th>Amount of noise introduced</th>
<th>% Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>No noise</td>
<td>500</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Low Noise</td>
<td>375</td>
<td>125</td>
<td>25.0</td>
</tr>
<tr>
<td>High Level of Noise</td>
<td>266</td>
<td>234</td>
<td>46.8</td>
</tr>
</tbody>
</table>

**COMPLEX EVENT DETECTION**

The generated event sequence file is then feed into CA-CED which is a tool for detecting the occurrences of CE given the appropriate event sequence file. CA-CED utilizes 4 similarity match techniques, which are (i) Sliding Window Detector (SWD), (ii) Weighted Sliding Window Detector (WSWD), (iii) Temporally Constrained Detector (TCD) and (iv) Multi-layered Event Detector for Generic Application (MEGA) for detecting the corresponding CE based on the pre-set event detection rule. All detectors employed in this research utilizes similarity match method in order to detect the CE by comparing the observed current and previous instance of Simple Events and the previous instance of CE with user-defined rule patterns (Mehdiyev et al. 2016). The SWD is a standard event detection algorithm while the WSWD, TCD and MEGA were relatively new event detection algorithm introduced by Saad (2017). The rule based template matching process will predict the CE occurrence by searching the desired sequence pattern in any one random event sequence. The comparison results are categorized into four (4) conditions as summarized in Table 2. Table 3 shows the description of five (5) types CE that exist in the dataset.

**TABLE 2. Event detection from rule template matching process**

<table>
<thead>
<tr>
<th>Actual Event</th>
<th>Detector Result</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Occurred</td>
<td>Not Detected</td>
<td>TN (True Negative)</td>
</tr>
<tr>
<td>Not Occurred</td>
<td>Detected</td>
<td>FP (False Positive)</td>
</tr>
<tr>
<td>Occurred</td>
<td>Not Detected</td>
<td>FN (False Negative)</td>
</tr>
<tr>
<td>Occurred</td>
<td>Detected</td>
<td>TP (True Positive)</td>
</tr>
</tbody>
</table>

**TABLE 3. Classification of simple events and complex event**

<table>
<thead>
<tr>
<th>Complex Event</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor_MovingRightFromR1 Actor_MovingRightFromR1⇒ Loitering</td>
<td></td>
</tr>
<tr>
<td>Actor_MovingLeftFromR1 Actor_MovingLeftFromR1⇒ Loitering</td>
<td></td>
</tr>
<tr>
<td>Actor_MovingRightFromR1 Actor_MovingRightFromR1⇒ Loitering</td>
<td></td>
</tr>
<tr>
<td>Actor_MovingLeftFromR1 Actor_MovingLeftFromR1⇒ Loitering</td>
<td></td>
</tr>
</tbody>
</table>

**RESULTS AND DISCUSSION**

A total of 500 annotated frame data were used to analyze the CE prediction performance of the four detectors. Four performance metrics which were: sensitivity, specificity, average accuracy and area under curve for ROC plot; have been used to evaluate the detector performance over CE detection. Sensitivity (True Positive Rate, TPR) and specificity (False Positive Rate, FPR) are statistical measures of a detector performance. Sensitivity measures how often the detector predicts the occurrence of a CE when the event occurs in real while specificity calculates how often the detector predicts
the occurrence of the CE when no event happens in real. A Confusion Matrix table was built to calculate the sensitivity and specificity of each detector for CE detection efficiency using varied confidence factor values of 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 and 0.0.

Table 4 shows the confusion matrix for the SuspiciousLoitering event class where the rule template matching procedure is conducted using confidence factor value of 1.0 with four (4) types of detector under the influence of three (3) types of noise. Table 5 summarizes the detection accuracy obtained. According to both tables, MEGA and TCD detectors have the higher average accuracy compared with another two detectors. Although WSWD has the highest TPR, it also has the highest FPR which shows the non-specificity of the detectors. On the other hand, MEGA and TCD detectors have very low FPR which shows the specificity of the detectors. Specificity is a very important criterion when evaluating detectors because an arbitrary detector will record the correct event happening at the wrong time.

The TPR and FPR values obtained from the confusion matrix table for each detector are used to plot the ROC curve so as to quantify the detector performance. Figure 5 (a), (b) and (c) demonstrate the ROC plot for four detectors on detection of event 501 that have been tested under three different noise values.

Results obtained showed that the TCD has greater ROC plot area in the no-noise and low noise conditions (area under ROC curve 0.5822 and 0.6146). Whereas, the MEGA detector has greater area under ROC of 0.6117 for the high-noise condition while maintaining an acceptable area under ROC curve for no-noise and low-noise conditions. Thus, we conclude that the MEGA detector is more suitable for use in the real-world condition, while maintaining reasonably acceptably performance under the no-noise and low-noise condition. Therefore, it is more suitable to be used in real ambient condition environment.

Figure 4 shows the results of video sequence for CE detection. This video frame is extracted from the CA-InViSS Lite system, where all monitored environment has been recorded. Each CE is represented by several SE occurred within the monitored areas.

FIGURE 4. Video sequence for CE detection
TABLE 4. Confusion matrix analysis for Suspicious Loitering event

<table>
<thead>
<tr>
<th>Detector</th>
<th>Perfect Noise - 0%</th>
<th>Intermediate Noise - 25%</th>
<th>Ambient Noise – 46.8%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Positive Rate (TPR)</td>
<td>False Positive Rate (FPR)</td>
<td>Average Accuracy</td>
</tr>
<tr>
<td>MEGA</td>
<td>0.4830</td>
<td>0.0230</td>
<td>0.9750</td>
</tr>
<tr>
<td>TCD</td>
<td>0.4860</td>
<td>0.0250</td>
<td>0.9740</td>
</tr>
<tr>
<td>SWD</td>
<td>0.5000</td>
<td>0.1780</td>
<td>0.8530</td>
</tr>
<tr>
<td>DSWD</td>
<td>0.1000</td>
<td>0.2510</td>
<td>0.7880</td>
</tr>
</tbody>
</table>

TABLE 5. Average detection accuracy for Suspicious Loitering event

<table>
<thead>
<tr>
<th>Main Parameter</th>
<th>Detector</th>
<th>Average Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEGA</td>
<td>97.20</td>
</tr>
<tr>
<td></td>
<td>TCD</td>
<td>97.50</td>
</tr>
<tr>
<td></td>
<td>SWD</td>
<td>89.70</td>
</tr>
<tr>
<td></td>
<td>DSWD</td>
<td>82.40</td>
</tr>
</tbody>
</table>

FIGURE 5. ROC plot with FPR & TPR as x & y axes, respectively for each detector for loitering event under a) no-noise condition (0% noise), (b) intermediate noise condition (25% noise) and (c) ambient noise condition (46.8% noise)
CONCLUSION

In conclusion, this study has successfully shown that its is possible to detect suspicious loitering event from CCTV feed using the SWD, WSWD, TCD and MEGA (average accuracies of 82.40, 89.70, 97.50, 97.20 respectively). The study also shows that MEGA gives the best overall performance, in particular when detection is done under ambient condition (area under ROC curve of 0.6117). The next step of the study is to implement the detection process in real-time into the Smart Surveillance System itself. Apart from that, the researchers would also look into developing algorithms to mine the annotated data and produce detection rule template automatically.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support provided by Malaysia’s Ministry of Higher Education and UniversitiKebangsaan Malaysia under Grant Nos. FRGS/1/2016/ICT02/UKM/02/7 and GUP-2017-124.

REFERENCES

*Rabiah Adawiyah Shahad, Mohd Faisal Ibrahim, Ezra Lim Kai Xian, Aini Hussain
Department of Electrical, Electronics & System Engineering,
Faculty of Engineering & Built Environment,
Universiti Kebangsaan Malaysia, Malaysia.

Mohamad Hanif Md Saad
Department of Mechanical & Materials Engineering,
Faculty of Engineering & Built Environment,
Universiti Kebangsaan Malaysia, Malaysia

*Corresponding author; email: rabiahshahad@siswa.ukm.edu.my

Received date: 6th December 2017
Accepted date: 18th February 2018
In Press date: 1st April 2018
Published date: 30th April 2018