Multimed Tools Appl (2014) 68:135–158
DOI 10.1007/s11042-012-1207-9

W^3-privacy: understanding *what, when*, and *where* inference channels in multi-camera surveillance video

Mukesh Saini · Pradeep K. Atrey · Sharad Mehrotra · Mohan Kankanhalli

Published online: 29 August 2012

Abstract Huge amounts of video are being recorded every day by surveillance systems. Since video is capable of recording and preserving an enormous amount of information which can be used in many applications, it is worth examining the degree of privacy loss that might occur due to public access to the recorded video. A fundamental requirement of privacy solutions is an understanding and analysis of the inference channels than can lead to a breach of privacy. Though inference channels and privacy risks are well studied in traditional data sharing applications (e.g., hospitals sharing patient records for data analysis), privacy assessments of video data have been limited to the direct identifiers such as people's faces in the video. Other important inference channels such as location (*Where*), time (*When*), and activities (*What*) are generally overlooked. In this paper we propose a privacy loss model that highlights and incorporates identity leakage through multiple inference channels that exist in a video due to *what*, *when*, and *where* information. We model the identity leakage and the sensitive information separately and combine them to calculate the privacy loss. The proposed identity leakage model is able to consolidate

M. Saini (✉) · M. Kankanhalli
School of Computing, National University of Singapore, Singapore, Singapore
e-mail: mksaini@comp.nus.edu.sg

M. Kankanhalli
e-mail: mohan@comp.nus.edu.sg

P. K. Atrey
Department of Applied Computer Science, The University of Winnipeg, Winnipeg, Canada
e-mail: p.atrey@uwinnipeg.ca

S. Mehrotra
Information and Computer Science Department, University of California, Irvine, CA, USA
e-mail: sharad@ics.uci.edu

the identity leakage through multiple events and multiple cameras. The experimental results are provided to demonstrate the proposed privacy analysis framework.

Keywords Privacy · Surveillance · Video · Modeling · Events · Anonymity

1 Introduction

Sensor traces have found vital applications in a variety of diverse scenarios, such as logistics, advanced driver assistance systems, medical care, public security, defense, aerospace, robotics, industrial production, precision agriculture [18], traffic monitoring, and policy making [2, 7]. However, the use of sensory data in these scenarios has often raised privacy concerns. The privacy concerns get more serious with video sensors. This is because people generally don't like their activities being recorded and watched by others. The main challenge here is to understand and analyze various inference channels that can result in a breach of privacy. While such inference channels are well studied in the context of traditional data sharing applications (e.g., a hospital releasing patient records and GPS based location aware services), it is challenging to understand inference channels embedded in semantically rich video.

In the past works on privacy preserving applications of video, it has been assumed that the identity leakage itself is equivalent to the privacy loss [5, 33]. We recognize that privacy loss occurs when an adversary is able to map an identity to the sensitive information present in the video, for example their habits, physique, companions, etc. The adversary can either be a human being with prior knowledge or an automated system with pattern information obtained through data mining and similar learning techniques [11]. There has been a great deal of work in privacy modeling in the field of video surveillance. In these works, privacy is modeled as a binary variable based on the presence of faces [5, 12, 19, 33] or silhouettes [15]. While removing the facial information is necessary for privacy preservation, it is not sufficient. Saini et al. [24] identified implicit identity leakage channels which exist even in the absence of the facial information. However, in this work the authors only considered privacy loss from a single camera video. The access to multiple camera videos may cause additional privacy loss in the following ways:

- The adversary can correlate persons in multiple video streams and observe more activities resulting in increased chances of identity leakage and privacy loss. For example, from a single camera generally we cannot infer what places a person visits or whom he meets over a period of time at different places. But this information can be easily extracted by correlating people over multiple camera videos.
- When an adversary identifies a person in a video, he can use this information to identity other persons in the video. Furthermore, this leakage could be propagated to other camera videos if the adversary is able to correlate people across multiple videos.

In the model proposed in [24], the activity information is measured as a binary variable for the whole video clip. It has completely ignored how many activities there are in the video. However, the chance of identity leakage increases with the amount of activity information. For example, it is easier to identify a person in a ten minute video full of activities rather than in a one hour video in which the activities effectively occur only for a short duration of one minute. The same is true for privacy loss; the chances of privacy loss generally increase with the number of activities in the video.

In this paper we analyze the implicit inference channels in the case of multi-camera surveillance videos and provide an enhanced model to measure the privacy loss due to *what*, *when*, and *where*, i.e. W^3-Privacy. In the proposed model, we first divide the video into a sequence of events and then analyze these events for privacy loss. Information extraction is applied to find different types of evidence *what, when* and *where*. The evidence information is used to measure the identity leakage that can occur due to each event. To incorporate the identity leakage due to event patterns, we prepare event lists for each person and fuse the identity leakages from all the events in the list using the anonymity based approach [28]. We separately measure the the amount of sensitive information as sensitivity index, and model privacy loss as a probabilistic product of identity leakage and the sensitivity index. While the proposed event based privacy model incorporates the breach in the current privacy methods, it also integrates seamlessly with currently growing research on event-based semantic representation of video [8, 10, 21]. The model measures the privacy at the semantic level in comparison to the earlier approaches which are mainly based on regions of interest (RoI) such as face and blob.

The main contributions of this paper are summarized as:

- A model that captures the increase in the identity leakage due to the simultaneous presence of multiple events.
- An assessment of privacy loss based on identity leakage and the associated sensitive information present in the video.
- Extending the model for privacy loss assessment to multi-camera surveillance videos.

The paper is laid out as follows. In Section 2, we describe the past works related to privacy loss assessment in video surveillance and discuss the novelty of the proposed work. Section 3 presents the proposed work. Various definitions used in the paper are provided in Section 3.1 and identity leakage is modeled in Section 3.2. A model to measure sensitivity of the video follows in Section 3.3 and privacy loss assessment in Section 3.4. We provide experimental results in Section 4, along with a discussion on the deployment of the model in real systems in Section 4.4. Finally we conclude the paper in Section 5.

2 Related work

In the past works, the privacy loss is often viewed as a discrete value. This set could be of size two (privacy is preserved or lost) [12, 15, 20, 23, 25, 29] or a fixed number [19, 33]. To the best of our knowledge, we are the first to model privacy loss as a continuous variable in the range [0,1] consolidating identity leakage through events and event patterns.

Privacy preservation has been studied by many researchers in the video surveillance community, leading to a number of data suppression techniques. In these techniques, the image regions occupied by humans are transformed partially or fully

Table 1 A summary of related work

Work	Implicit channels	Multi-camera	Sensitive information	Modeling binary/continuous	Consideration of events
Boyl et al. [4]	No	No	No	Binary	No
Senior et al. [25]	No	No	No	Binary	No
Moncrieff et al. [19]	No	No	No	Fixed levels	No
Fidaleo et al. [12]	No	No	No	Binary	No
Wickramasuriya et al. [33]	No	No	No	Fixed levels	No
Koshimizu et al. [15]	No	No	No	Binary	No
Spindler et al. [27]	No	No	No	Fixed levels	No
Thuraisingham et al. [29]	No	No	No	Binary	No
Carrillo et al. [5]	No	No	No	Binary	No
Paruchuri et al. [20]	No	No	No	Binary	No
Qureshi et al. [23]	No	No	No	Binary	No
Saini et al. [24]	Yes	No	No	Continuous	Single
Proposed model	Yes	Yes	Yes	Continuous	Multiple

to hide the identity information. For instance, [12] proposed to use a filter to detect and remove the facial information before saving it onto the server until the person's behavior is considered suspicious. Similar strategies are adopted in works [4, 5, 25] to preserve privacy in video, i.e. if a person's face is obscured in the video, the privacy is considered preserved. In another set of works, the whole body is replaced by a solid color, estimated background, bar, dot, edges, border, silhouette, or mosaic [6]. While these works don't model privacy loss explicitly, they assume that hiding bodily cues (such as faces) is enough for privacy preservation. They have overlooked other implicit inference channels associated with *what* (activity), *where* (location where the video is recorded) and *when* (time when the video is recorded). An adversary can observe the behavior, look at the places visited and use prior knowledge to infer identity information.

Table 1 presents a summary of existing works and shows how the proposed work is novel in comparison. This summary has been provided from the following aspects: whether implicit inference channels are considered; whether privacy loss is modeled for surveillance video from multiple cameras; whether the notion of sensitive information has been used in privacy loss computation; whether privacy loss is determined as a binary or continuous value; and whether privacy loss is determined based on single or multiple events in the video. It is clearly shown in the table that the proposed work is novel in many aspects.

3 Proposed work

3.1 Definitions

For the sake of brevity, let us first define the key terms used in the paper.

Definition 1 (Event) The event definition has been adopted from [3]: 'Event is a physical reality that consists of one or more living or non-living real world objects (who) having one or more attributes (of type) being involved in one or more activities (what) at a location (where) over a period of time (when)'.

Definition 2 (Identity leakage) The probability with which an adversary can identify an individual in the video.

Definition 3 (Anonymity) The size of the group of people who can not be distinguished from each other with the information available in the video.

Definition 4 (Sensitivity index) This is a measure of sensitive information in the video for which an individual feels a privacy violation would occur if made available to public.

Definition 5 (Privacy loss) This is defined as the probability that an adversary will be able to gain sensitive information about an individual depicted in the video.

Definition 6 (Implicit channels) Except facial information, all other means of identity leakage are called implicit channels of identity leakage.

Definition 7 (Evidence) This is the information which is extracted from the video data. The facial information is termed as "who" evidence, any activity related information is termed as "what" evidence, temporal information is said to be "when" evidence, and spatial information as "where" evidence. These evidence types have been identified as main aspects of a generic event model [32].

3.2 Identity leakage

In the proposed framework, the privacy loss is determined based on the identity leakage and association of the identity with the sensitive information present in the video. This is illustrated in Fig. 1.

The aim of an adversary is to establish a relation between an identity and sensitive information present in the video. Hence, the first step in privacy modeling is to determine the extent of identity leakage. Let us start with the analysis of the human

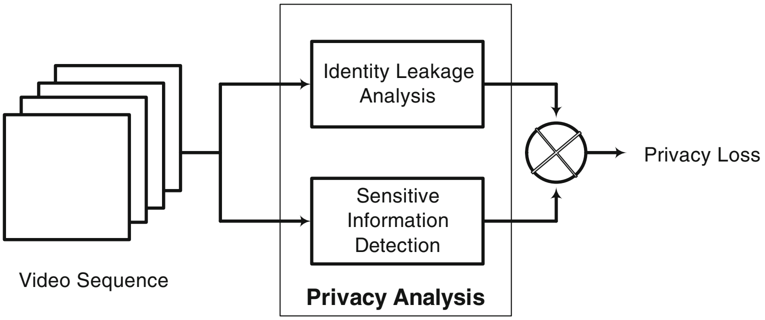


Fig. 1 The process of assessment of privacy loss of individuals in the video

Table 2 Different idiosyncrasies which human beings use in order to recognize other people

Evidence	Idiosyncrasy	Example
who	Face	The face is considered to have features which distinguish people accurately.
what	Clothes	Depending on their taste, people repeat a particular style of clothes.
	Gait	Many people have a particular way of moving their body parts.
	Who else they meet	A person meeting a professor very frequently is probably one of his students.
	Social behavior	This includes gender specific, culture specific, and religion specific behavior.
when	Timing of actions	What time they have lunch, what time they go to the office etc.
where	Spatial information	A person inside a particular food stall is most likely the owner.

recognition system. We recognize people generally by their name, face, and habits etc. Table 2 enumerates usual idiosyncrasies used by human beings to recognize fellow humans. In our formalization of identity leakage, we model 'face' as *who* evidence, 'gait' and 'social behavior' as *what* evidence, 'time' as *when* evidence, and spatial information as *where* evidence [24]. Other aspects such as 'associated objects' and 'who else they meet' can be considered by determining the number of objects present in the video.

We measure the identity leakage through the degree of anonymity. An identity leakage with κ-anonymity means that the size of the smallest group with which the adversary can associate the individual's identity is κ [28]. With the detection of each idiosyncrasy, we are able to associate the identity of the individual to a subgroup of people (the default initial group is the world's total population). For example, when it is detected that the place is 'Smart Lab, NUS', through prior knowledge we can relate the identity of the individuals in the video to the group of people who visit 'Smart Lab'. Furthermore, if the time is also detected as evening, we can use the knowledge that only half of them are expected to be there, reducing the association group by half. Hence, the identity leakage depends on prior knowledge of the adversary and corresponding observations from the video. We model the knowledge of an adversary as a rule based expert system [14]. In its knowledge base, an expert system contains facts and beliefs learned over time, which can be used to interpret the observations (in our case its purpose is to infer identity from given evidence). The structure of the knowledge base and its application will be discussed later. Note that a significant amount of work is being done on knowledge modeling in the interdisciplinary fields of Natural Language Processing, Information Retrieval, Machine Learning, and Knowledge Representation and Reasoning [31]. Systems are being developed that can learn these rules automatically [10, 11] and make inferences, causing privacy loss.

Figure 2 shows the overall framework for identity leakage calculation. Events are detected from the video of each camera and analyzed to enumerate the number of targets. Consequently, an event list is constructed for each target. In the figure, $T_1, T_2, \ldots, T_m$ denote m targets and $E_1, E_2, \ldots, E_m$ denote the corresponding event

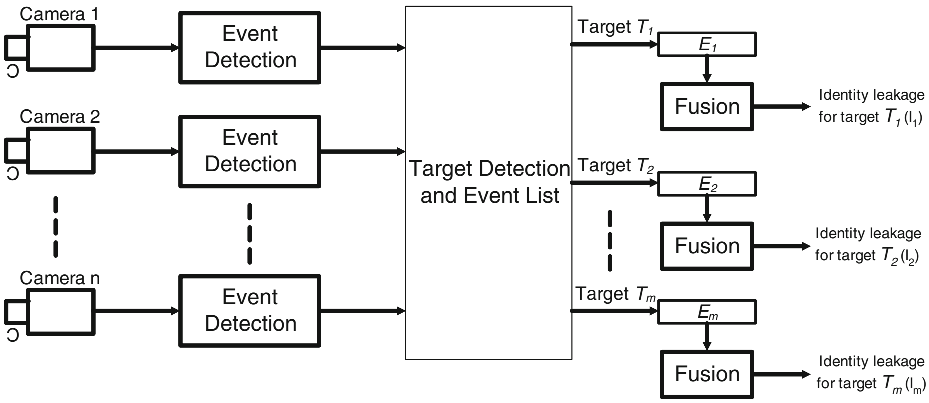

Camera 1
Camera 2
Camera n
Event Detection
Event Detection
Event Detection
Target Detection and Event List
Target T_1
Target T_2
Target T_m
E_1
E_2
E_m
Fusion
Fusion
Fusion
Identity leakage for target T_1 (I_1)
Identity leakage for target T_2 (I_2)
Identity leakage for target T_m (I_m)

Fig. 2 The framework for **Identity Leakage Analysis**. Here, the term target is used to denote individuals in the video

lists. The event list contains all the events in which a particular target is detected. Overall identity leakage I_i for the target T_i is calculated by using anonymity based fusion [28] of the information from the corresponding events in the event list E_i. The anonymity is calculated for each target appearing in the detected events using the adversary's knowledge base.

3.2.1 Video segmentation

The input video from multiple cameras is segmented into events. This segmentation can be performed based on various criteria. For example, one event can be the video segment between two consecutive background frames encompassing non background frames with motion and activity [17]. Alternatively, video can be segmented based on the number of people [26]. The proposed framework is independent of how we segment the video.

3.2.2 Evidence detection

After segmenting the video into events, we detect evidence from each event. This is done by analyzing the information present in the video. If the information is sufficient to recognize the place, we consider that the *where* evidence is detected. Similarly, if the event contains time information, it leads to *when* evidence. The *what* evidence is always present if the event involves one or more persons; which is true by definition (c.f. Section 3.1).

3.2.3 Proposition generation

The event related knowledge of an adversary can be represented using propositional logic statements. This knowledge can be learned using machine learning techniques or it can be an expert knowledge base related to the application scenario. Every statement consists of a premise and conclusion. A premise is a proposition that is used as the foundation for drawing a conclusions. In our case, each premise proposition statement consists of information about an event. An event consists of following attributes: *when*, *where*, and *what* (we exclude the *who* because the face is assumed obscured in the video). Hence, a premise statement is represented by a 4-tuple $\mathcal{P}(t_s, t_e, act, loc)$; where t_s and t_e are the start and end time of an event respectively (when), *act* is the activity (what), and *loc* is the location (where). For example, a premise statement $\mathcal{P}$(25-Oct-2010:10:35, 25-Oct-2010:11:10, “working”, “smart lab”) means “a group of people was working in the smart laboratory from 10:35 h to 11:10 h on 25-Oct-2010”. This premise leads to the corresponding conclusion $\mathcal{C}(grp)$, where *grp* is the associated group of people. A knowledge base consists of pairs of premise and conclusion statements. A knowledge base entry $\mathcal{P} \Rightarrow \mathcal{C}$ denotes that if a premise $\mathcal{P}$ is true, it leads to conclusion $\mathcal{C}$. For example, $\mathcal{P}$ (25-Oct-2010:10:35, 25-Oct-2010:11:10, “working”, “smart lab”) $\Rightarrow \mathcal{C}$(G1) would mean that “The group of people who are working in the smart laboratory from 10:35 h to 11:10 h on 25-Oct-2010 are G1”. Note that the absence of any attribute in the tuple is represented by a null symbol ‘ϕ’. For instance, a premise statement $\mathcal{P}$ (ϕ, ϕ,“dancing”, ϕ) denotes “a group of people was involved in dancing activity in the scene”. This premise leads to the corresponding conclusion $\mathcal{C}$(G2) “The group of people in the video who are involved in the dancing activity are G2”.

Using the above propositional statements, we can build propositions for each type of idiosyncrasy listed in Table 2. For example, identity leakage through clothes can be represented as $\mathcal{P}(t_s$, t_e, “kurta”, “office”) to distinguish people who wear kurtas in an office, social behavior related identity leakage can be represented as $\mathcal{P}(t_s$, t_e, “praying”, “temple”) which can map to the people of a particular religion who go to temple to pray, the proposition for gait related identity leakage can be constructed as $\mathcal{P}(t_s$, t_e, “hand in pocket”, “temple”), and companion related identity leakage as $\mathcal{P}(t_s, t_e$, “with Mukesh”, “temple”).

The event contains the information that can be learned by the adversary about any individual. Every event is a potential source of identity leakage. The identity leakage can take place in two ways:

- Through individual events in isolation, i.e. every event has *what, when* and *where* information which can be used to associate a person’s identity to a particular group of people.
- Through a spatiotemporal sequence of events which might map to an identity revealing patterns present in the knowledge base. The matching patterns further restrict the identity to a subgroup of people.

Although all statements of the knowledge base conclude in an association group, they differ for both the cases discussed above. The statements used for identity leakage from events generally have propositions that are generated by only that event, whereas for a later case the premise may consist of multiple propositions derived from multiple events which might be from multiple cameras. We will calculate both types of identity leakages (due to individual events and event patterns) and combine them to find the overall identity leakage.

3.2.4 Identity leakage from individual events

The propositions generated in the previous section contain the event information. However, the identity leakage only occurs if similar propositions are also present in the adversary’s knowledge base. Hence, for each proposition, we find mapping in the knowledge base. If an appropriate mapping is found, we add the corresponding associated group in the set of mapped groups $\mathcal{G}$. Let $\mathcal{G}^e$ be the resulting association group due to event information. It is calculated as the intersection of all the groups in $\mathcal{G}$ as follows:

$$\mathcal{G}^e = \cap\{G \mid G \in \mathcal{G}\} \tag{1}$$

For example, if the knowledge base consists of propositions $\mathcal{P}_1$ to $\mathcal{P}_{10}$ with the corresponding association groups G1 to G10; and the event under consideration generates propositions $\mathcal{P}_2$ and $\mathcal{P}_8$, then the set of mapped groups $\mathcal{G} = \{G2, G8\}$ and the resulting association group $\mathcal{G}^e = G2 \cap G8$. The above equation implies that by using the information present in the event, we are able to associate the identity of the person seen in the video to a group of people $\mathcal{G}^e$.

3.2.5 Identity leakage through multiple event patterns

In the previous section we modeled the identity leakage by considering the events in isolation. However, the adversary may track the person over multiple events and multiple cameras which allows the adversary to exploit the knowledge of event patterns to further reduce the anonymity of identity. In order to model the identity leakage through event patterns, events are analyzed to detect the total number of targets present in the video. A separate event list is created for each target as shown in Fig. 2. The target association across events is done based on similarity of appearances such as height, clothes etc. This is because the adversary can obtain an event sequence only if the person looks similar across cameras.

Once we build the event list for all the targets, we fuse the identity leakage for each target using the information obtained from its associated event list. Let $\mathcal{G}^e_{ij}$ be the association group for the jth event in event list E_i, which corresponds to target T_i, and is calculated using (1).

To understand how the knowledge of patterns helps in identity leakage, let us consider the following example. Suppose the adversary knows that G1 = {A1, A2, ..., A10} people are expected at site 1, and G2 = {A1, A2, B1, B2, ..., B6} people are expected at site 2. If A1 appears at site 1 and site 2 in two separate events (from separate cameras), the corresponding anonymities are 10 for event 1 and 8 for event 2. This results in an anonymity of 8 for A1 (i.e. the minimum of the two values), if events are considered in isolation. However, even without facial information the adversary can identify that the person detected at site 1 as well as site 2 is same through visual similarity. The adversary has the pattern information that only A1 and A2 are seen at both sites which results in a reduced anonymity of 2. The pattern information (only A1 and A2 are seen at both sites) is embedded in G1 and G2 and it can be easily obtained by intersecting G1 and G2. Hence, the combined association group $\mathcal{G}'_i$ for target T_i is calculated as the intersection of all the association groups of corresponding events in its event list E_i (See (2)).

$$\mathcal{G}'_i = \mathcal{G}^e_{i1} \cap \mathcal{G}^e_{i2} \cap \ldots \mathcal{G}^e_{in_e} \tag{2}$$

where n_e is the number of matching propositions for all the events in E_i.

In the discussion above, it is assumed that the adversary has the knowledge of all populations in their entirety. In practice the adversary may not have knowledge of G1 and G2 but may only know that if someone is seen at both site 1 and site 2, it is either A1 or A2, which can be stored in the knowledge base using the following statement:

$$\mathcal{P}(\phi, \phi, \phi, Site1) \wedge \mathcal{P}(\phi, \phi, \phi, Site2) \Rightarrow \mathcal{C}(\text{A1}, \text{A2}) \tag{3}$$

Now, if there are two events generating propositions $\mathcal{P}(\phi, \phi, \phi, \text{Site1})$ and $\mathcal{P}(\phi, \phi, \phi, \text{Site2})$, they will not cause any identity leakage individually as they do not have any mapping statements. However, when found to be related to the same target, both events can be mapped together to the pattern statement discussed above causing additional identity leakage. Let $\mathcal{G}^p_{i1}, \mathcal{G}^p_{i2}$... be the association groups for the matching pattern statements. The overall association group is now calculated by intersecting all the association groups corresponding to events ($\mathcal{G}^e_{ij}$) and event patterns ($\mathcal{G}^p_{ij}$):

$$\mathcal{G}_i = \left(\mathcal{G}^e_{i1} \cap \mathcal{G}^e_{i2} \cap \ldots \mathcal{G}^e_{in_e}\right) \cap \left(\mathcal{G}^p_{i1} \cap \mathcal{G}^p_{i2} \cap \ldots \mathcal{G}^p_{in_p}\right) \tag{4}$$

In the above equation, n_p is the number of matching patterns for all the events in E_i.

The anonymity, κ_i, of target T_i is calculated as the size of the overall association group $\mathcal{G}_i$:

$$\kappa_i = |\mathcal{G}_i| \tag{5}$$

Finally, the identity leakage, I_i, for target T_i is the inverse of the anonymity by definition and is computed as:

$$I_i = \frac{1}{\kappa_i} \tag{6}$$

If all the events belong to one person, that person should be common among all the association groups. Nevertheless, there can be multiple people satisfying the same set of propositions. For example, there can be multiple people who come to the laboratory at night and do similar activities. Therefore, the identity leakage is generally less than one.

3.3 Sensitivity index

Privacy loss and identity leakage are two separate phenomena. Mere identity leakage does not always lead to privacy loss. For example, a video which only shows the full frontal face reveals the identity of the person quite accurately. However, if no other information can be learned from the video (activity, place, time, etc.), people generally don't feel that their privacy is being compromised, whereas, if the video also shows which place the person is visiting or whom the person is meeting, it might be a privacy loss for some individuals. Similar situations can be found in the statistical data publication where well structured data records of individuals are published after the removal of the direct identifiers [13]. There, the privacy loss occurs when an adversary is able to map the identity to the sensitive information stored in the sensitive information fields of the published data records. For example, medical data records might contain disease names as sensitive information.

Video generally contains an enormous amount of information which might qualify as *sensitive*. Which information is sensitive and which is normal, depends on the individuals and may vary from person to person [16]. Yet Table 3 enumerates commonly found video attributes which are considered sensitive. These attributes can be categorized as one of the evidence types *what, when*, and *where*, and the

Table 3 Commonly identified sensitive information

Sensitive attribute	Example
Activity	Adjusting clothing when alone.
Spatial information	Generally we do not want strangers to know which places we visit.
Time	Some people mind when others associate their activities with timing patterns.
Gesture	People make strange gestures while they are alone and do not want others to watch that.
Clothes	Many teens wear clothes which they do not want their parents to see.
Physique	People with an atypical physique may be sensitive to that, e.g. height.
Habits	Most people have some personal idiosyncratic sensitive habits like twiddling fingers under stress.
Companion information	Some people do not want everyone to know with whom they associate.
Associated objects	What we carry with us.

identity leakage through these implicit channels can provide a basis to determine the privacy loss. Further discussion on state of the art techniques to determine these attributes is provided in Section 4.4.

We assume that the information in the video consists of a set of attributes and that some of these represent the sensitive information. Let $A = \{a_1, a_2, ...a_l\}$ be the set of attributes that can potentially be sensitive information. Let $\mathcal{W}$ be the priority vector defined as:

$$\mathcal{W} = \{w_k \mid k \in [1, l], w_k \geq 0, w_1 + w_2, ... + w_l = 1\} \tag{7}$$

The elements of the vector are weights (w_k) which are set by the individuals seen in the video and they reflect their priority of the corresponding attribute as sensitive information. While the priority vector could be different for each individual, we recommend the calculation of a representative priority vector for a group of people seen at video recording site. For each target we detect the sensitive attributes in all of the events in the corresponding event list and build a sensitivity matrix as follows:

$$S = \{s_{ki} \mid k \in [1, l], i \in [1, m]\} \tag{8}$$

Each column is related to one target, and each row to one sensitive attribute. The elements of the array are calculated as follows:

$$s_{ik} = \begin{cases} 1 \text{ if } k\text{th attribute is detected for target } T_i; \\ 0 \text{ otherwise.} \end{cases} \tag{9}$$

The sensitivity index for each ith target can be calculated as follows:

$$\Psi_i = \sum_{k=1}^{l} w_k s_{ik}. \tag{10}$$

The equation reflects that any information in the video only adds to the privacy loss if it is also sensitive to the individuals in the video. In the above discussion we assume that the *sensitivity* of an attribute remains unchanged over time. However, an attribute which is not sensitive today, may become sensitive at some later point in time. This is the limitation of all existing privacy solutions and remains so in our proposed model as well. It is very difficult to capture the evolving sense of privacy loss.

3.4 Privacy loss

If we remove the identity information completely, the video cannot cause privacy loss to any individual, no matter how much sensitive information the video contains. This is because the sensitive information cannot be associated with anyone. Similarly, if there is nothing sensitive in the video, it generally does not cause privacy loss even when people are identified. In both the cases, the resulting privacy loss is zero. Hence, the privacy loss γ_i of ith target can be calculated as a product of the identity leakage (I_i) and the sensitivity index Ψ_i, i.e.

$$\gamma_i = I_i \Psi_i \tag{11}$$

where γ_i is the total privacy loss of ith individual. In order to reduce the privacy loss, we need to minimize both identity leakage and sensitive information. Interestingly, even if one of them is close to zero, we can keep privacy loss to a minimum.

3.5 Remarks

The proposed model measures the privacy loss of the individuals that are usual inhabitants living in the vicinity of the video recording site. The framework takes a conservative approach and models the privacy loss that would occur when no random people appear at the site. To understand this point, let us assume that there is a hypothetical model that considers random people as well in determining the association groups. If r is the number of additional random people seen at the site, $\mathcal{G}$ is the association group according to the proposed framework, and $\mathcal{G}^r$ is the association group considering these r random people as well; it can be easily proven that $\forall r \geq 0, \mathcal{G} \subset \mathcal{G}^r$. In the limiting case we have:

$$\lim_{r \to 0} \frac{|\mathcal{G}^r|}{|\mathcal{G}|} = 1. \tag{12}$$

Since $\gamma \propto 1/|\mathcal{G}|$ (see (5), (6), and (11)), the privacy loss computed using our model will always be higher than the actual privacy loss for any nonzero r. The proposed privacy model considers the worst case scenario when no random people show up and provides an upper bound of the privacy loss for the given video and adversary knowledge base.

This paper examines an anonymity based approach to ensure privacy in a video, however, it remains to be explored how a differential privacy mechanism [9], which is a completely new direction and claims to provide a more accurate measure of privacy, can be used in the W3-privacy framework.

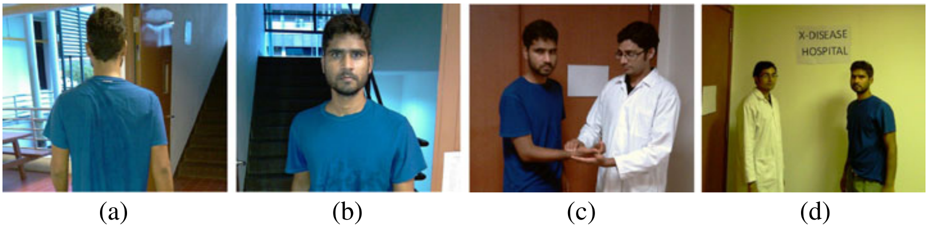
4 Experimental results

To demonstrate the utility of the proposed model, we conduct three experiments. In the first experiment we highlight the difference between identity leakage and privacy loss. In the second, the effect of multiple events of the identity leakage is shown. Finally in the third experiment, we show how the proposed framework is used to calculate privacy loss in the case of a multi-camera surveillance video.

4.1 Experiment 1: identity leakage vs privacy loss

The distinction between identity leakage and privacy loss is demonstrated using the following experiment. We consider a case where a person is sick and visits a hospital. He does not want his friends to know his disease and the doctor with whom he is consulting, likely because the doctor's specialization might reveal the disease. Here, there are two sensitive attributes $A = \{a_1, a_2\}$ where a_1 is the companion and a_2 is the location. The corresponding priority weights for sensitive information are given as $W = \{w_1, w_2\}$ where $w_1 = 0.5$, and $w_2 = 0.5$. He goes to the hospital and the surveillance camera records four separate images of the person. For this example we assume that each image is representative of one event. The four images are shown in Fig. 3. The sensitivity matrix is a column vector since we are analyzing the privacy loss of only the patient, which is denoted as $S = \{s_1, s_2\}^T$. The values of s_1 and s_2 are different for different images which are determined below.

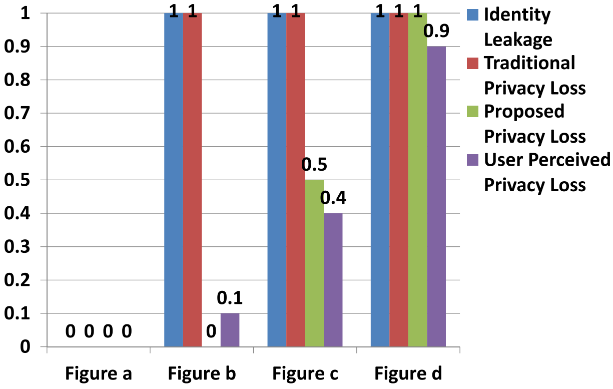
In Fig. 3a, we cannot see the person's face, hence the privacy loss predicted by traditional models is 0. The picture also does not have any other information which can be used for implicit inference channel, hence the proposed model also gives zero identity loss as well as privacy loss. In Fig. 3b we can see the person's face implying a privacy loss of 1 with traditional models of privacy. However, since no sensitive information is available in this image, s_1 and s_2 are taken as 0, which results in zero privacy loss using the proposed model (11). Figure 3c clearly has companion information, making $s_2 = 1$. This results in a privacy loss of $\gamma = 0.5 \times 1 = 0.5$. Finally, from Fig. 3d we can deduce the exact disease through hospital name. Hence the privacy loss $\gamma = 1 \times (0.5 \times 1 + 0.5 \times 1) = 1$. Figure 4 shows the results of the experiment. To get the user perceived privacy loss, five students aged between 20 and 30 were explained the situation and were asked to rate the privacy loss for each image from 0 to 1. The users were explained the scenario that a person has a sensitive disease and he is visiting a doctor in a hospital. If they are the sick person in the picture, how would they rate the privacy loss for each of the four pictures.



(a) (b) (c) (d)

Fig. 3 Four pictures taken by surveillance cameras placed around a hospital

Fig. 4 Identity vs privacy



There are two implications of the results. Firstly, if the video does not contain any sensitive information for the person, we do not need to hide the identity information; the video will not cause any privacy loss. Alternatively, if the identity cannot be inferred from the video, the video can be released with all the sensitive information, although this is generally not possible as sensitive information itself can cause identity leakage through other implicit inference channels.

4.2 Experiment 2: event based identity leakage

The goal of this experiment is to highlight the effect of multiple events on identity leakage. We use a video clip from a single camera recorded in a laboratory scenario. The knowledge base in this case has statements shown in Table 4. The video, half an hour in length, consists of seven events described in Table 5 with three targets involved. Here, SL means 'Smart Lab' and for clarity we mention the events as $C_1, C_2, \ldots,$ etc., which later form the event lists E_i. The representative images from four of them are shown in Fig. 5. In this video clip, all three types of evidence are detected as follows: *what* = walking (WK), running (RN), discussion (DC), *where* = smart lab (SL), *when* = evening (EV). Since all the events had the same starting and ending time (EV), in the propositions we mention both the start and end time using a single EV.

Since the proposition generation depends on the targets, we describe it for individual targets separately. In Table 6, the first column shows the events in which the target is detected. The second column shows the proposition generated by that event. In the third column we write the association group according to the mapping proposition in the knowledge base shown in Table 4. We calculate the final association group by calculating the intersection of all groups due to individual events. Table 6 shows the results for T_1 and T_2. Target T_3 only appears in one event and generates the proposition $\mathcal{P}$(EV, WK, SL) which maps to statement 3, giving $\mathcal{G}_3$ = G3. In case the event generated proposition does not have mapping

Table 4 Knowledge base for experiment 2

1. $\mathcal{P}(\phi, \phi, \text{SL}) \Rightarrow$ G1(A1–10)
2. $\mathcal{P}(\phi, \text{DC}, \text{SL}) \Rightarrow$ G2(A1–2, A4–7)
3. $\mathcal{P}(\text{EV}, \phi, \text{SL}) \Rightarrow$ G3(A3–4, A7, A9)
4. $\mathcal{P}(\phi, \text{RN}, \text{SL}) \Rightarrow$ G4(A5–7, A9–10)

Table 5 Event description of the video for experiment 2

Event	Description	Event	Description	Event	Description
C_1	T_1, WK	C_4	T_1, WK	C_7	T_2, RN
C_2	T_1, WK	C_5	T_1, T_2, DC	–	–
C_3	T_2, WK	C_6	T_3, WK	–	–

in the knowledge base, the association group is assumed to be the universal group (UV) which is superset of all other groups. Note that if only individual events are considered, the association groups would be the smallest in the list, which is incidentally the same for all three targets, i.e. 4. Figure 6 compares anonymities calculated using both models.

It is interesting to observe that a unity identity leakage of a target can increase identity leakage of other targets. For example, in this experiment we conclude that the anonymity of T_1 is anonymous between A4 and A7 and T_2 is known to be A7. Since we know that A7 is T_2, we can conclude that T_1 is A4. The exact identities of T_1 and T_2 also reduce the anonymity of T_3 to two (A3 or A9). The ground truth is A4, A7, A9.

4.3 Experiment 3: privacy loss from multiple cameras

If the adversary has access to multiple camera videos where the same person is spotted at multiple places, the adversary can use the knowledge of spatiotemporal patterns to infer the identity. In this experiment we demonstrate two main contributions: (1) The effect of multiple cameras on identity leakage and (2) The privacy loss assessment in a multi-camera scenario. For this experiment we recorded video at four places in the department building: (1) Department Entrance (DE) (2) Audio Lab (AL) (3) Staff Club (SC) and (4) Canteen (IC) Fig. 7. A total of 40 people are expected in the department (A1–10, B1–10, C1–10, D1–10). However, the adversary does not have knowledge about all of them. The adversary's knowledge is limited to the propositional logic statements given in Table 7.

Six actors created a series of events at these sites. Table 8 provides the description of all the events captured at these sites. These actors were involved in one of the following activities: discussion, walking, and running. For the calculation of anonymity, we created event lists for each target separately. The time was detected as evening in all the cameras except camera 1 at the department entrance.

The event lists for the targets and generated propositions are given in Table 9. Similar to the previous experiment, the first column shows the events where the target was detected. The second column shows the proposition generated by the

(a) (b) (c) (d)

Fig. 5 Representative images from four events of the video recoded in the smart lab

Table 6 Event lists and identity leakage for individual targets

E_1 (Target T_1)			E_2 (Target T_2)		
C_1	$\mathcal{P}$(EV, WK, SL)	G3	C_3	$\mathcal{P}$(EV, WK, SL)	G3
C_2	$\mathcal{P}$(EV, WK, SL)	G3	C_7	$\mathcal{P}$(EV, RN, SL)	G4
C_4	$\mathcal{P}$(EV, DC, SL)	G2	C_4	$\mathcal{P}$(EV, DC, SL)	G2
	$\mathcal{P}$(EV, WK, SL)	G3		$\mathcal{P}$(EV, WK, SL)	G3
C_5	$\mathcal{P}$(EV, WK, SL)	G2	–	–	–
$\mathcal{G}_1 = G2 \cap G3 = (A7, A9)$			$\mathcal{G}_2 = G2 \cap G3 \cap G4 = (A7)$		

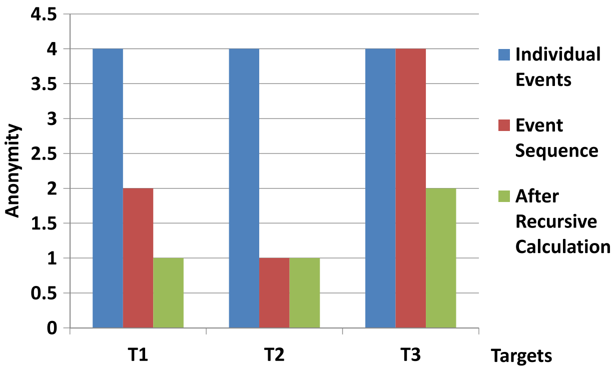
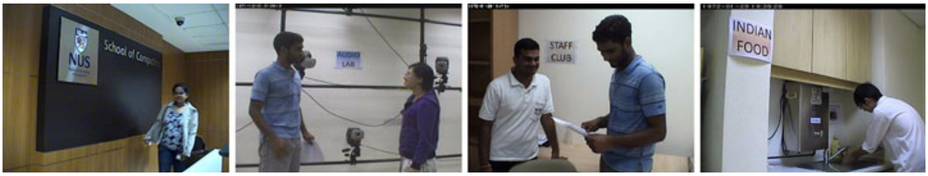


Fig. 6 The anonymity when we consider events in isolation and event sequences. The *third bar* shows the results of recursive identity leakage



(a) DE (b) AL (c) SC (d) IC

Fig. 7 Representative images from four cameras: **a** Department Entrance, **b** Audio Lab, **c** Staff Club, **d** Canteen

Table 7 Knowledge base for experiment 3

Statements for individual events

1. $\mathcal{P}(\phi, \phi, \text{SC}) \Rightarrow$ G1(B1–10, C1–2, A1–3)
2. $\mathcal{P}(\phi, \phi, \text{AL}) \Rightarrow$ G2(A1–10, B1–5, D1–2)
3. $\mathcal{P}(\phi, \phi, \text{IC}) \Rightarrow$ G3(A1–5, B1–5, C8–10)
4. $\mathcal{P}(\phi, \text{DC}, \phi) \Rightarrow$ G4(A1–3, B1–4, D5–8)
5. $\mathcal{P}(\text{EV}, \phi, \text{SC}) \Rightarrow$ G5(A1–2, B1–5, C1–2)
6. $\mathcal{P}(\text{EV}, \phi, \text{AL}) \Rightarrow$ G6(A1–4, B1–3, D1–2)
7. $\mathcal{P}(\text{EV}, \phi, \text{IC}) \Rightarrow$ G7(A1–5, B1–3, C8)

Statements for multi-event patterns

8. $\mathcal{P}(\phi, \phi, \text{DE}) \wedge \mathcal{P}(\phi, \phi, \text{IC}) \Rightarrow$ G8(A1, A8, B1, B8)
9. $\mathcal{P}(\phi, \phi, \text{DE}) \wedge \mathcal{P}(\phi, \phi, \text{SC}) \Rightarrow$ G9(A1, B1–3, C1–3)
10. $\mathcal{P}(\phi, \phi, \text{DE}) \wedge \mathcal{P}(\phi, \phi, \text{AL}) \Rightarrow$ G10(A1–6, B1–6, D1)

Table 8 Description of events captured by cameras

Event	Description	Event	Description	Event	Description
Camera 1 - Department Entrance (DE)					
C_{11}	T_1, WK	C_{16}	T_4, WK	C_{112}	T_4, WK
C_{12}	T_1, WK	C_{17}	T_6, WK	C_{113}	T_2, RN
C_{13}	T_5, WK	C_{18}	T_5, WK	C_{114}	T_2, WK
C_{14}	T_2, WK	C_{19}	T_2, WK	–	–
C_{15}	T_3, WK	C_{110}	T_3, WK	–	–
Camera 2 - Audio Lab (AL)					
C_{21}	T_1, T_2, DC	C_{23}	T_1, WK	C_{25}	T_1, RN
C_{22}	T_6, RN	C_{24}	T_2, WK	–	–
Camera 3 - Staff Club (SC)					
C_{31}	T_4, WK	C_{35}	T_5, WK	C_{39}	T_5, WK
C_{32}	T_3, WK	C_{36}	T_5, WK	C_{310}	T_3, WK
C_{33}	T_4, WK	C_{37}	T_4, WK		
C_{34}	T_3, T_1, DC	C_{38}	T_5, WK		
Camera 4 - Canteen (IC)					
C_{41}	T_1, WK	C_{43}	T_4, WK	C_{45}	T_4, WK
C_{42}	T_1, WK	C_{44}	T_1, WK		

event. In third column we write the association group derived from the knowledge base by proposition mapping. An event proposition may have multiple matches in the knowledge base, in which case we list all the groups in the third column.

In order to calculate privacy loss, we need to determine the sensitive information matrix. In this experiment we have chosen the sensitive attributes to be: (1) Companion (2) Running activity (3) Height (4) Clothes. The priorities for the attributes are as follows: $\mathcal{W} = \{0.45, 0.30, 0.15, 0.10\}$. The weights have been determined based on common notions of privacy. People are more sensitive to their company than their clothes or height. Similarly, they might not feel comfortable being watched while running. The sensitivity matrix can be easily derived from event descriptions as follows:

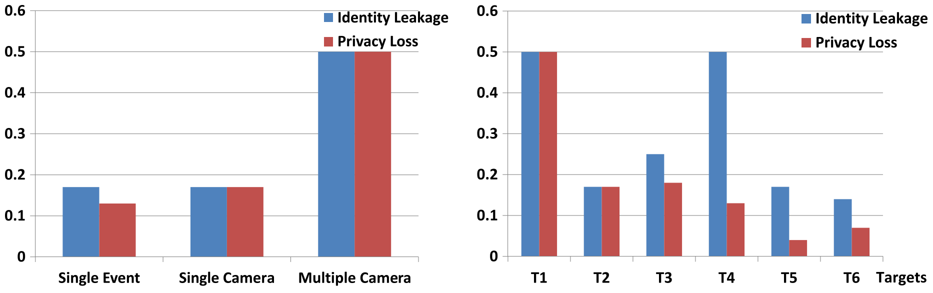
$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \tag{13}$$

Since $I_i = 1/|\mathcal{G}_i|$, the identity leakage vector is calculated as $I = \{0.50, 0.17, 0.25, 0.5, 0.17, 0.14\}$. With these values of W, S, and I, we can calculate the privacy loss. Figure 8a shows the resulting identity leakage and privacy loss in three scenarios of identity leakage: (1) Individual events (2) Patterns among single camera events and (3) Patterns among multiple camera events. It can be seen that when multiple camera video is available and adversary has knowledge of patterns, the identity leakage and privacy loss increases.

Similarly, Fig. 8b shows the identity leakage and privacy loss for all the targets measured using the proposed framework for multi-camera video. The identity leakage for T_1 is highest because T_1 was seen at all four sites and was involved in all the activities: walking, running and discussion. T_2 and T_4 appear in the same number of events, however, T_4 appears in events from multiple cameras hence its identity leakage is higher than T_2. T_1 and T_4 have the same identity leakage, yet the privacy

Table 9 Event lists and identity leakage for targets

E_1 (Target T_1)					
C_{11}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{41}	$\mathcal{P}(\text{EV}, \text{WK}, \text{IC})$	G7
C_{12}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{42}	$\mathcal{P}(\text{EV}, \text{WK}, \text{IC})$	G7
C_{112}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{44}	$\mathcal{P}(\text{EV}, \text{WK}, \text{IC})$	G7
C_{21}	$\mathcal{P}(\text{EV}, \text{DC}, \text{AL})$	G4, G6	Pattern	8	G8
C_{23}	$\mathcal{P}(\text{EV}, \text{WK}, \text{AL})$	G6	Pattern	9	G9
C_{25}	$\mathcal{P}(\text{EV}, \text{RN}, \text{AL})$	G6	Pattern	10	G10
C_{34}	$\mathcal{P}(\text{EV}, \text{DC}, \text{SC})$	G4, G5			
$\mathcal{G}_1 = \text{G4} \cap \text{G5...G10} = (\text{A1}, \text{B1})$					
E_2 (Target T_2)					
C_{14}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{21}	$\mathcal{P}(\text{EV}, \text{DC}, \text{AL})$	G4,G6
C_{19}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{24}	$\mathcal{P}(\text{EV}, \text{WK}, \text{AL})$	G6
C_{113}	$\mathcal{P}(\phi, \text{RN}, \text{DE})$	UV	Pattern	10	G10
C_{114}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV			
$\mathcal{G}_2 = \text{G4} \cap \text{G6} \cap \text{G10} = (\text{A1–3}, \text{B1–3})$					
E_3 (Target T_3)					
C_{15}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{34}	$\mathcal{P}(\text{EV}, \text{DC}, \text{SC})$	G4, G5
C_{110}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{310}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5
C_{32}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5	Pattern	9	G9
$\mathcal{G}_3 = \text{G4} \cap \text{G5} \cap \text{G9} = (\text{A1}, \text{B1}-3)$					
E_4 (Target T_4)					
C_{16}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{43}	$\mathcal{P}(\text{EV}, \text{WK}, \text{IC})$	G7
C_{31}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5	C_{45}	$\mathcal{P}(\text{EV}, \text{WK}, \text{IC})$	G7
C_{33}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5	Pattern	8	G8
C_{37}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5	Pattern	9	G9
$\mathcal{G}_4 = \text{G5} \cap \text{G7} \cap \text{G8} \cap \text{G9} = (\text{A1}, \text{B1})$					
E_5 (Target T_5)					
C_{13}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	C_{38}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5
C_{18}	$\mathcal{P}(\text{EV}, \text{WK}, \text{DE})$	UV	C_{39}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5
C_{35}	$\mathcal{P}(\text{EV}, \text{WK}, \text{SC})$	G5	Pattern	9	G9
$\mathcal{G}_5 = \text{G5} \cap \text{G9} = (\text{A1}, \text{B1}-3, \text{C1}-2)$					
E_1 (Target T_6)					
C_{17}	$\mathcal{P}(\phi, \text{WK}, \text{DE})$	UV	Pattern	10	G10
C_{22}	$\mathcal{P}(\text{EV}, \text{RN}, \text{AL})$	G6	–	–	–
$\mathcal{G}_6 = \text{G6} \cap \text{G10} = (\text{A1}-4, \text{B1}-3, \text{D1})$					



(a) (b)

Fig. 8 **a** Identity leakage and privacy loss for T_1. **b** Identity leakage and privacy loss for all targets loss of T_1 is still higher than T_4. This shows the effect of the sensitive attributes on privacy loss. For T_1, all sensitive attributes are detected whereas for T_4, only two out of four attributes are detected.

4.4 Discussion and practical considerations

This paper highlights the privacy breach that exists in the current privacy preserving methods which consider only facial and appearance based cues. Although the current technology is not robust enough to decipher the event information accurately, a human observer can definitely detect *what, when,* and *where* information from multi-camera video which may lead to privacy loss. In experiments it is demonstrated that even when the bodily cues are absent, in extreme cases the adversary can identify the individuals with a small value of anonymity.

The success of any privacy preserving method depends on the automated detection techniques and there is a whole community of researchers working on improving these detectors [8, 10, 21]. Recent surveys on human activity detection techniques [1, 22, 30] describe how current research is progressing toward the detection of different gestures, actions, interactions and group activities. While gestures include individual body part movements such as “stretching an arm” and “raising a leg”; actions are composed of a sequence of gestures such as “walking”, “jogging”, “hand shaking”, “pushing” “pool diving”, “boxing”, “kissing”, “hitting”, “opening a cabinet”, “picking up an object”, “jumping”, “bending”, and “waving”. Similarly, interactions and group activities involve multiple people e.g. “meeting”, “marching”, “fighting”, “presenting”, “discussing”, “taking a break” etc. Furthermore, there have been a number of works on localization using video, which can automatically detect the presence of *where* information [34–36].

We recognize that the extent to which we can block these privacy leakage channels is limited by the accuracy of the methods for automatic object tracking and event detection. However, we can follow a conservative approach for event detection in order to preserve privacy (but at the cost of utility). By “conservative approach”, we mean that we can lower the detection thresholds. Although this may result in a large number of false detections; by adopting this approach we may over-do some suppression operations. This approach can be flexible in the sense that the thresholds can be raised as progress is made in event detection research. In this paper we restricted our focus to expose the various channels (other than the human face) that can cause privacy leakage, which we believe is an important first step toward future privacy aware multi-camera systems.

The acceptable value of privacy loss can be determined by considering the sensitivity of the video recording site and application context. For example, in a very sensitive application, such as defense, even a small value of privacy loss may not be acceptable, while in other normal applications like shopping mall surveillance, a relatively higher value of privacy loss could be acceptable.

5 Conclusions & future work

Implicit inference channels of *what*, *when*, and *where* can cause significant privacy loss when an adversary gets access to multiple-camera surveillance videos. The privacy loss through these inference channels is modeled as W^3-Privacy. The privacy loss measured by the proposed model is closer to the user perceived privacy loss than earlier models. Furthermore, privacy loss only occurs when sensitive information and identity leakage co-exist. Therefore, any of these can be separately controlled to minimize privacy loss. The proposed model can be configured for any adversarial knowledge and the sensitive attributes of the individuals, making it flexible and applicable in diverse scenarios. For example, in a surveillance scenario, the occupants of the surveilled premises can provide the sensitive attributes and the person who has access to this surveillance video can be considered as an adversary.

Following are the conclusions derived from this work:

1. The proposed model is able to consolidate the privacy loss through *what*, *when*, and *where* inference channels. Compared to earlier models, the proposed model calculates privacy loss that is closer to the user perceived privacy loss.
2. Event and activity patterns across multiple cameras lead to privacy loss which does not result from a single camera. Therefore, a recurring behavior at multiple camera sites can increase the risk of privacy loss for the usual occupants.
3. Identity leakage of an individual can also affect the identity leakage of other individuals in the video. When the identity leakage of a person is unity, the anonymity of all other people showing similar patterns reduces by 1.
4. The proposed model measures the upper bound of privacy loss by considering the worst case scenario when only usual occupants are seen(i.e., no random people appear).

The privacy assessment model proposed in this paper is necessarily the first and a very important step toward privacy protection of the people in multi-camera videos, and this work sets the directions for future research, i.e. to investigate methods to reduce the privacy loss with minimal loss of utility of videos. We hope to inspire further research in data transformation techniques, which will retain in the video just enough information required for intended application (e.g. surveillance), without compromising the privacy of the people under surveillance.

References

1. Aggarwal J, Cai Q (1997) Human motion analysis: a review. In: Proc. of IEEE nonrigid and articulated motion workshop, pp 90–102
2. Akyildiz I, Su W, Sankarasubramaniam Y, Cayirci E (2002) A survey on sensor networks. IEEE Commun Mag 40(8):102–114
3. Atrey P, Kankanhalli M, Jain R (2006) Information assimilation framework for event detection in multimedia surveillance systems. Multimedia Syst 12(3):239–253
4. Boyle M, Edwards C, Greenberg S (2000) The effects of filtered video on awareness and privacy. In: The ACM conference on computer supported cooperative work, pp 1–10
5. Carrillo P, Kalva H, Magliveras S (2008) Compression independent object encryption for ensuring privacy in video surveillance. In: IEEE international conference on multimedia and expo, pp 273–276
6. Chinomi K, Nitta N, Ito Y, Babaguchi N (2008) Prisurv: privacy protected video surveillance system using adaptive visual abstraction. In: Proceedings of the international conference on advances in multimedia modeling, pp 144–154

7. Chong C, Kumar S (2003) Sensor networks: evolution, opportunities, and challenges. Proc IEEE 91(8):1247–1256
8. Doulamis A, van Gool L, Nixon M, Varvarigou T, Doulamis N (2008) First ACM international workshop on analysis and retrieval of events, actions and workflows in video streams. In: ACM international conference on multimedia, pp 1147–1148
9. Dwork C (2006) Differential privacy. In: International colloquium on automata, languages and programming, pp 1–12
10. Fernández C, Baiget P, Roca FX, Gonzílez J (2011) Determining the best suited semantic events for cognitive surveillance. Expert Syst Appl 38(4):4068–4079
11. Ferrucci D (2010) Build watson: an overview of deepqa for the jeopardy! challenge. In: International conference on parallel architectures and compilation techniques, pp 1–2
12. Fidaleo D, Nguyen H, Trivedi M (2004) The networked sensor tapestry (nest): a privacy enhanced software architecture for interactive analysis of data in video-sensor networks. In: ACM iternational workshop on video surveill ance & sensor networks, pp 46–53
13. Fung B, Wang K, Chen R, Yu P (2010) Privacy-preserving data publishing: a survey on recent developments. In: ACM computing surveys, vol 42
14. Hayes-Roth F, Waterman D, Lenat D (1984) Building expert systems. Addison-Wesley, Reading, MA
15. Koshimizu T, Toriyama T, Babaguchi N (2006) Factors on the sense of privacy in video surveillance. In: ACM workshop on continuous archival and retrival of personal experences, pp 35–44
16. Langheinrich M (2001) Privacy by design - principles of privacy-aware ubiquitous systems. In: International conference on ubiquitous computing. Springer, pp 273–291
17. Lu Y, Ga W, Wu F (2002) Automatic video segmentation using a novel background model. In: The IEEE international symposium on circuits and systems, pp 807–810
18. McBratney A, Whelan B, Ancev T, Bouma J (2005) Future directions of precision agriculture. Precis Agric 6(1):7–23
19. Moncrieff S, Venkatesh S, West G (2008) Dynamic privacy assessment in a smart house environment using multimodal sensing. ACM Trans Multimed Comput Commun Appl 5(2): 1–29
20. Paruchuri JK, Cheung S, Hail MW (2009) Video data hiding for managing privacy information in surveillance systems. In: SPIE newsroom
21. Piciarelli C, Foresti G (2011) Surveillance-oriented event detection in video streams. In: IEEE intelligent systems, pp 32–41
22. Poppe R (2010) A survey on vision-based human action recognition. Image Vis Comput 28(6):976–990
23. Qureshi FZ (2009) Object-video streams for preserving privacy in video surveillance. In: International conference on advanced video and signal based surveillance, pp 442–447
24. Saini M, Atrey P, Mehrotra S, Emmanuel S, Kankanhalli M (2010) Privacy modeling in video data publication. In: IEEE international conference on multimedia and expo, pp 60–65
25. Senior A, Pankanti S, Hampapur A, Brown L, Tian YL, Ekin A, Connell J, Shu CF, Lu M (2005) Enabling video rivacy through computer vision. IEEE Secur Priv 3(3):50–57
26. Septian H, Tao J, Tan YP (2006) People counting by video segmentation and tracking. In: International conference on control, automation, robotics and vision, pp 1–4
27. Spindler T, Wartmann C, Hovestadt L, Roth D, Van Gool L, Steffen A (2006) Privacy in video surveilled areas. In: The ACM international conference on privacy, security and trust, pp 1–10
28. Sweeney L (2002) k-anonymity: a model for protecting privacy. Int J Uncertain Fuzziness Knowl-Based Syst 10(5):557–570
29. Thuraisingham B, Lavee G, Bertino E, Fan J, Khan L (2006) Access control, confidentiality and privacy for video surveillance databases. In: ACM symposium on access control models and technologies, pp 1–10
30. Turaga P, Chellappa R, Subrahmanian V, Udrea O (2008) Machine recognition of human activities: a survey. IEEE Trans Circuits Syst Video Technol 18(11):1473–1488
31. Van Harmelen F, Lifschitz V, Porter B (2008) Handbook of knowledge representation. Elsevier Science Ltd
32. Westermann U, Jain R (2007) Toward a common event model for multimedia applications. IEEE Multimed 14(1):19–29
33. Wickramasuriya J, Datt M, Mehrotra S, Venkatasubramanian N (2004) Privacy protecting data collection in media spaces. In: International conference on multimedia, pp 48–55
34. Wiles R, Hirvonen D, Hsu S, Kumar R, Lehman W, Matei B, Zhao W (2001) Video georegistration: algorithm and quantitative evaluation. In: Proc. of IEEE international conference on computer vision, vol 2, pp 343–350

35. Zhu Z, Oskiper T, Samarasekera S, Kumar R, Sawhney H (2007) Ten-fold improvement in visual odometry using landmark matching. In: Proc. IEEE international conference on computer vision, pp 1–8
36. Zhu Z, Oskiper T, Samarasekera S, Kumar R, Sawhney H (2008) Real-time global localization with a pre-built visual landmark database. In: Proc. of IEEE conference on computer vision and pattern recognition, pp 1–8

Mukesh Saini is a research fellow at School of Computing, National University of Singapore (NUS). He obtained his Master of Technology (M. Tech.) in Electronics Design and Technology from Indian Institute of Science (IISc), Bangalore, in 2006 and his PhD in Computer Science from School of Computing, National University of Singapore, Singapore in 2012 respectively. During his PhD stint, he visited University of Winnipeg for six months (Sep 2009 to Feb 2010) under the Canadian Commonwealth Exchange Program. His PhD dissertation is on the cutting-edge topic "Privacy-aware multimedia surveillance". His other research interests include generic architecture for observation systems and system level performance evaluation of multimedia systems. He has played roles of Reviewer, TPC member, Tutorial organizer, and Panelist for various reputed conferences and journals.

Pradeep K. Atrey is an Associate Professor and a Senator at the University of Winnipeg, Canada. He received his PhD in Computer Science from National University of Singapore, M.S. in Software Systems and B.Tech. in Computer Science and Engineering from India. He was a Postdoctoral Researcher at the Multimedia Communications Research Laboratory, University of Ottawa, Canada.

His current research interests are in the area of Multimedia Computing with a focus on Multimedia Surveillance and Privacy, Image/Video Security, and Web. He has authored/co-authored over 55 research articles at reputed ACM, IEEE, and Springer journals and conferences. Dr. Atrey is on the editorial board of ETRI Journal and Journal of Convergence (Web and Multimedia). He is actively involved in his research community and he has been associated with over 20 international conferences in various roles such as General Chair, Program Chair, Publicity Chair, Web Chair, and TPC Member. Dr. Atrey was recipient of the ETRI Journal Best Reviewer Award (2009) and the University of Winnipeg Merit Award for Exceptional Performance (2010). He was also recognized as "'ICME 2011 - Quality Reviewer".

Sharad Mehrotra is a Professor in the School of Information and Computer Science at University of California, Irvine (UCI) and Director of the Center for Emergency Response Technologies (CERT) at UCI. Mehrotra's research expertise is in data management and distributed systems areas in which he has made many pioneering contributions. Two such contributions include the concept of "database as a service" and "use of information retrieval techniques, particularly relevance feedback, in multimedia search". Mehrotra is a recipient of numerous best paper nominations and awards including SIGMOD Best Paper award (2001), Best of VLDB submissions (1994), and best paper award in DASFAA (2004). Mehrotra's current research focuses on building sentient spaces using multimodal sensors, data privacy, and data quality. Mehrotra's recent research, particularly, in the context of CERT has focused on situational awareness from multimodal input including conversational speech data. Many of his research contributions have been incorporated into software artifacts which are now in use at various first responder partner sites. He holds a patent on privacy protection of data collection in pervasive environments.

Mohan Kankanhalli is a Professor at the Department of Computer Science of the National University of Singapore. He is also the Associate Provost for Graduate Education at NUS. Before that, he was the Vice-Dean for Academic Affairs and Graduate Studies at the NUS School of Computing during 2008–2010 and Vice-Dean for Research during 2001–2007. Mohan obtained his BTech (Eletrical Eng.) from the Indian Institute of Technology, Kharagpur, in 1986 and his MS and PhD (Computer Systems Eng.) from the Rensselaer Polytechnic Institute in 1998 and 1990, respectively. He was a researcher at the Institute of Systems Science at NUS during 1990–1997. He then became a faculty member at the Department of Electrical Engineering of the Indian Institute of Science, Bangalore during 1997–1998 after which he joined NUS again. He visited the University of California at Berkeley during Jan–Jun 2004. He is actively involved in the Multimedia Systems community and he is currently the Director of Conferences for ACM SIG Multimedia. He is on the editorial boards of several journals including the ACM Transactions on Multimedia Computing, Communications, and Applications, Springer Multimedia Systems Journal, Pattern Recognition Journal and Multimedia Tools & Applications. His current research interests are in Multimedia Systems (content processing, retrieval) and Multimedia Security (surveillance, digital rights management and privacy).