PROFILE DATA DEPLOYING STRATEGY FOR EFFICIENT COOPERATIVE REASONING ON ONTOLOGY

Soomi Yang

Department of Information Engineering, The University of Suwon, Gyeonggido, Korea
smyang@suwon.ac.kr

ABSTRACT

This paper introduces the profile data management scheme such as data structuring, interchanging and data deployment. An overwhelming amount of multimedia data is generated by surveillance devices such as smart cameras. Profile data contained in them if properly structured and is integrated, can induce a useful context information. This work builds a hierarchical profile data deploy structure and import related multimedia data to annotate rich data arriving from multiple sensor streams, in this case smart cameras. The annotation process provides an impetus to the improvement of knowledge over time. Proactive deploying provides the main concepts and properties to model a hierarchical area data structure which can span a university campus or an apartment or a city. We also define management policies to compare their performance for the wide area surveillance specifically.

Profile Data Deploying, Reasoning based on Ontology, Wide Area Surveillance System, Context Inference, Smart Camera Network

1. INTRODUCTION

The popularity of smart cameras providing many intelligent functions grows with the advance of electronic technologies and Internet applications. It explores the convergence of the caching and streaming technologies for multimedia data. For the surveillance of the large area, agents built in networked RFID sensors, CCTVs and smart cameras should collaborate through integration of profile data. Profile data includes audio, video, recognized feature data, ontology data and others. Most such transfers would take place with the streaming data passing through one or more of the agents caching the streaming data as it passes through. Effective caching techniques will be critical for the successful deployment of streaming multimedia data over the surveillance network. This should be obvious, because of the huge latencies involved, and the requirements of real time play out. This paper describes a framework for the working of such a distributed caching and streaming system. Distributed agents receiving heterogeneous data from various sources have autonomy, collaborate with each other, and do ontology reasoning based on distributed knowledge bases. A framework for the integration of profile data supplied by a set of agents is designed. The agents are interconnected through a peer network. And non-leaf administrative agents in a geographic area hierarchy or in an administration hierarchy build a hierarchical structure according to their right hierarchy. This framework guarantees the consistency and expressivity of the ontology used in the data integration process. In the process of reasoning each agent may process the consolidated data for the distributed and autonomous reasoning, which is scalable and efficient[1], helps security persons by giving appropriate decision or prediction based on huge ontology data about situation it gathers.

The distribution of demands for profile data items is often skewed, and the surveillance devices have different capabilities and data formats. These can lead to poor data communication and dropped messages. In this paper, we propose a proactive caching scheme for efficiency and

DOI : 10.5121/ijdps.2011.2602

interoperability. Implementation is also going on into our distributed surveillance network environment.

The rest of the paper is organized as follows, Section 2 surveys related work of cooperative inference schemes. Our proposed modelling framework is explained in Section 3. Section 4 describes the adaptive profile data management technique. In Section 5, simulation results are presented and the performance is evaluated. We also show our implementation results. Finally, Section 6 concludes with an outline of our future work.

2. RELATED WORK

There have been a lot of cooperative data deploying schemes proposed with analogical peers. Numerous studies have been carried out on the caching and replication techniques in distributed environments. Especially, [2] dealt with multimedia data which is the most interested data format for the surveillance environment. In [3], a cooperative caching framework is introduced and claimed to be effective to data availability. In [4], a replica allocation method and clustering in distributed networks are introduced to improve data accessibility in a mobile communication environment. Another cooperative data management scheme for similar peers is described in [5], which combines the P2P communication technology with a conventional mobile system. [6] tries to allocate fair buffer to regulate a varying traffic, and [7] suggests a proactive network management solution. However, previous works were mostly based on the ground that the distributed nodes in the network have the same characteristics, and the data transmitted is standardized and intermittent. Our distributed surveillance environment consisted of various data source devices, multimedia data is generated continuously and analysis should be done in real time.

To ease merging of heterogeneous data, effort for the standardization for physical security is done by ONVIF(Open Network Video Interface Forum)[8] and PSIA(Physical Security Interoperability Alliance)[9]. They define, recommend, and promote standards for IP-based security products. Our implemented system tried to meet the requirements of industry by providing functions recommended by standard organizations. There are few of surveillance systems adopting ontology-driven technologies. [10] introduces artificial intelligence techniques only for the interpretation of objects. [11] uses ontology but does not build agents for web of data. Furthermore existing wide area surveillance system are closed system and do not provide surveillance information as a public web services which is suggested in the standard.

To address these challenges, we advocate an adaptive hierarchical profile data deploy framework in which heterogeneous devices with different transmission ranges, latency and other factors co-exist in the distributed surveillance network with a proactive data management scheme to provide efficient public web services.

3. PROPOSED SURVEILLANCE FRAMEWORK

3.1. Hierarchical Infrastructure

The architecture consists of a hierarchy of agents, which contains a number of non-leaf node administrative agents and leaf node agents with sensors, as shown in Figure 1. It consists of national agents at the top, some regional agents at the middle level, and many local agents at the second lowest level. At the lowest level, data source agents such as agents built in smart camera make leaf nodes. Though data source agents can be directly connected to higher level agents, most of data source agents are connected to the local agents. Meanwhile, the users are directly connected to any level of agents. Each agent will be allowed to source the profile data into the caches ahead of any client requests. Thus, the agent may not be on-line all the time but its content can always be assumed available and hosted in one or more of the caching agents for the integrated reasoning. This proactive caching also considers load balancing aspect. The rectangles mean region servers and the ovals mean data source servers. The data source server

can be a single camera node or sensor node. Possibly it can be a back end server for several cameras or sensor nodes. The lowest level server is called 'level 1 server'. Level 1 region servers receive data from data source servers in their domain and manage them. Rarely upper level region servers can have data source server of their own.

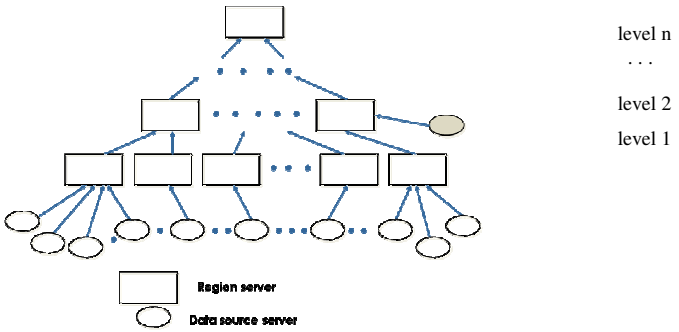

Figure 1. Hierarchical infrastructure of region servers

The servers form a graph structure. They can communicate each other freely within access control permission to perform their own intelligent distributed inference based on their own ontology knowledge base.

3.2. System Framework

We are designing and deploying a system for the wide area surveillance system covering a province area containing about 10 cities. The design of architecture for the surveillance of the area consists of an integrated framework of networked RFID sensors, CCTVs, and smart cameras. The surveillance device's purpose is not only to take pictures, record videos and log location data, but also to analyze a scene and report items and activities of interest to the user. Also it is required to take actions based on prediction by ontology reasoning.

Various data, such as video, feature data including biometrics, event alarms, originate from many kinds of input devices. The input devices can be deployed as fixed devices or on moving robots for dangerous situations like floods, earthquake, etc. Although, agents are distributed by geographic area hierarchy, they have their own knowledge base including ontology data about situation, inference engine and communication protocol between input devices for queries. Deploying profile data in distributed smart camera surveillance network has usually a hierarchical architecture on top of the hierarchical administration organization. That is, each hierarchical level of administrative agents is configured to operate in the same level, which is usually placed at the access point connecting to the higher level of agents. These data, often referred to as institutional, regional, and national data depending on the hierarchy of the administrative agents operating them, is integrated with each other so as to infer and deliver efficiently the requested result to the users.

With a profile data distribution, even small agents, with the help of higher level agents, can achieve very high hit rate locally, thus reducing significantly the bandwidth requirements of the network connection and the latency perceived by the users. To pursue this performance gain, the deploying of profile data in administrative agents and proper distribution structure is needed for the better hit rates achievable by the keeping profile data only in local data store with peer network.

Despite this potential gain of the administrative agent distribution, there still remain many technical and managerial problems to be solved before it could fulfill the gain in full. The

administrative agent's data store might be more heavily loaded as the number of data increases faster and their update intervals go shorter and shorter. The situation becomes worse as more data are replaced by multimedia data. This then cause the middle level agents to suffer a long period of time before they contain most of the data of agents in the surveillance camera network, degrading the whole system performance significantly.

This paper assumes a proactive, rather than reactive, deploying and attempts to explore an optimal profile data management strategy and service policy for surveillance networks. By proactive deploying, it means that the middle level agents cooperate with the master in exchanging necessary information on the data size, the request rate and so on enabling the master to fetch and broadcast the data before a lower level agents requests it. To achieve this goal, the surveillance network is modeled as a multi-layered distribution network in the next section. And the system performances like hit rate, disk space, bandwidth gain of the agents and the leaf node source servers, bandwidth for the surveillance network, and latency experienced by the users are derived.

4. ADAPTIVE PROFILE DATA MANAGEMENT

As shown in Figure 1, the profile data including multimedia data and feature data will be distributed over the region servers and data source servers. The data source server will need to carry out the indexing and retrieval of the information distributed across the servers in an efficient manner. To aid the task, we could use a data structure containing only the information related to the requested objects.

In the hierarchical profile data structure, agents cooperate with each other in such a way that the request not satisfied at the lower level is forwarded to a higher level in the structure hierarchy until it satisfied. If the request is not satisfied at any level of agents, it is forwarded to the original data source server. As contrasted to this hierarchical request chaining, forwarding of request at the same level operates quite a different scheme on top of the surveillance network architecture. It forms another overlay network.

With the proactive deploying, which we assume in this paper, the master agent fetches data from the data source server based on the exchanged information and broadcasts them to the subscribing agents before any institutional agent requests them. This service has no hierarchy among the agents and all agents maintain the same set of data for some designated suspect. With this service, any missed request at the local agent is directly forwarded to the original data source server.

In this section, we specify some features of the forwarding profile data and present analytical models for evaluating the system performances.

Define D={1,2,…,|D|} as the whole set of profile data, and assume that the data in D are ranked in the order of their suspiciousness or importance. From the well-known results that the probability of the *i*-th data is collected is Zipf distributed[12], we have the probability function,

$$f(i) = \frac{\sigma}{i^{\alpha}}, i = 1,2,...,|D| \tag{1}$$

where $\sigma = (\sum_{i \in D} 1/i^{\alpha})^{-1}$ and α is a constant close to 1.

Let n_k be the total number of connected agents and λ_k be the data request rate, respectively. It is natural to assume that λ_k is proportional to n_k.

If agents do not differ from each other in their preferences for data, the request rate of agent *k* for the data *i*, denoted by λ_{ki} , can be written as

$$\lambda_{ki} = \lambda_k f(i) = \beta n_k \cdot \frac{\sigma}{i^{\alpha}} \tag{2}$$

Let D_F be the set of profile data required. Also, Let S_i be the size of the *i*-th data. Then, any subscribing agent k needs to have an institutional storage with the required disk space δ_k , which amounts to,

$$\delta_k = \sum_{i \in D_F} S_i \tag{3}$$

Note that every agent operates the same set of profile data for some designated suspect, the expression for δ_k being the same and independent *k*.

With the same preference distributions and the same set of profile data, all the subscribing agents have the same hit rate as a distributed surveillance environment, multimedia data and feature data are produced continuously. To maintain the freshness and effectiveness of the data, we should profile data adaptively. We use the term 'weight' of data to describe its relative importance as compared to the other data as proposed in [13]. The higher the weight, the lower is the probability of the clip being replaced. We also use a policy based on the size of the objects, in which the weight is proportional to the inverse of the size of the data for the network bandwidth usage efficiency. Therefore the weight *w* is computed as following, where *F* is the number of times the data is accessed, *S* is the size of the data and *R* is the time since the last access for the data. The three exponents *f* , *r* and *s* are weighting factor.

$$\begin{aligned} w &= F^f S^s R^r \\ &= {\lambda_{ki}}^f {\delta_k}^s {\mu_i}^r \\ &= (\beta n_k \cdot \frac{\sigma}{i^{\alpha}})^f \cdot (\sum_{i \in D_F} S_i)^s \cdot {\mu_i}^r \end{aligned} \tag{4}$$

where μ_i is the update interval. The value for *f* should be a positive number, meaning that more frequently accessed data is more likely to be found. The value of *s* can be should be a negative number for the efficient use of network bandwidth, such that more small data is more likely to be stored. The value of *r* should be a negative number, meaning that more recent data is

more likely to be stored. If the recentness is more important than the frequency, the absolute value of the exponent r should be greater than that of the exponent *f*.

5. PERFORMANCE EVALUATION

5.1. Simulation

For each given request rate β, we can inspect the effects to the data weight diversity. Figure 2 shows the expected data weight by data rank *i* and data update interval μ_i. With high ranked data and small update interval, the data weight *w* grows high. Furthermore we can see the more high data weight for the higher request rate β.

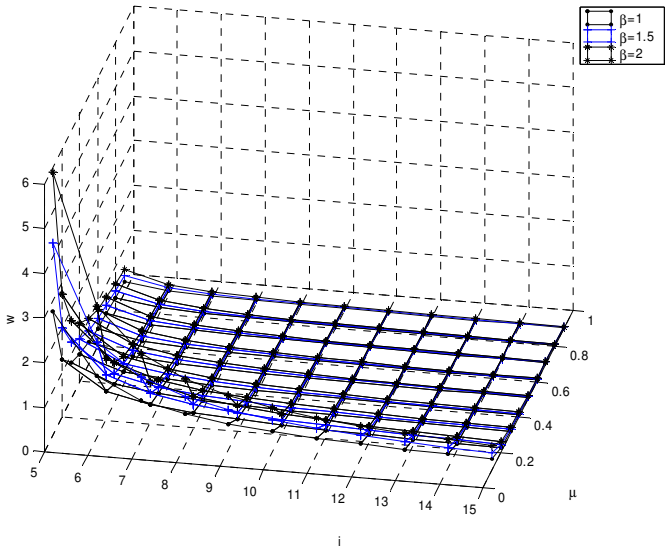

Figure 2. The data weight diversity by request rate

Figure 3 shows the data miss rate diversity with different data weight. The difference between the two graphs comes from the difference of the data weight that was seen in Fig. 2. To evaluate the effect to the proactive data caching, we compare the data miss rate for the different weight threshold for the data replacement using similar method in [14]. The data miss rate is calculated relative to the case without proactive data caching. For the adaptive deploying technique, we can raise the threshold value for the proactive caching. When we increase the threshold value, the data miss rate *Q* decreases as shown in Figure 3.

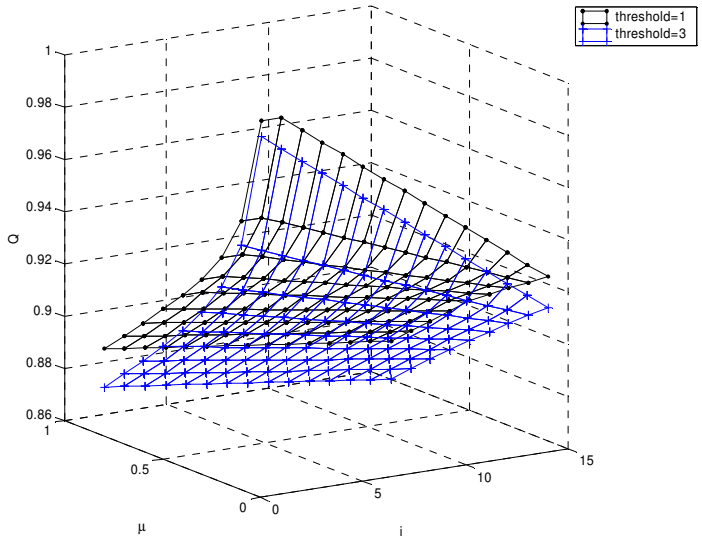

Figure 3. The data miss rate diversity with different weight threshold

5.2. Implementation

We implemented our cooperative framework for integrating ontologies with an adaptive deploying technique into our distributed surveillance system. Figure 4 and 5 show user interfaces. Web-based user interface shown in Figure 4 lists biometric data recognized on the right frame. On the left frame includes google map and marks for peculiar events. The event lists pop up is the results of the integrated inference. When we click each of the event items, it pops up related information. The subject tracking through the communication between the independent data source servers is possible as shown in Fig. 5. They request and inquire profile data for a combined inference through defined queries on the peer to peer computing framework. In Figure 5, we show the grids used for the cooperative inference. Mobile user interface is also implemented to see the real time or stored video and for smart camera control

Figure 4. Web interface with google map

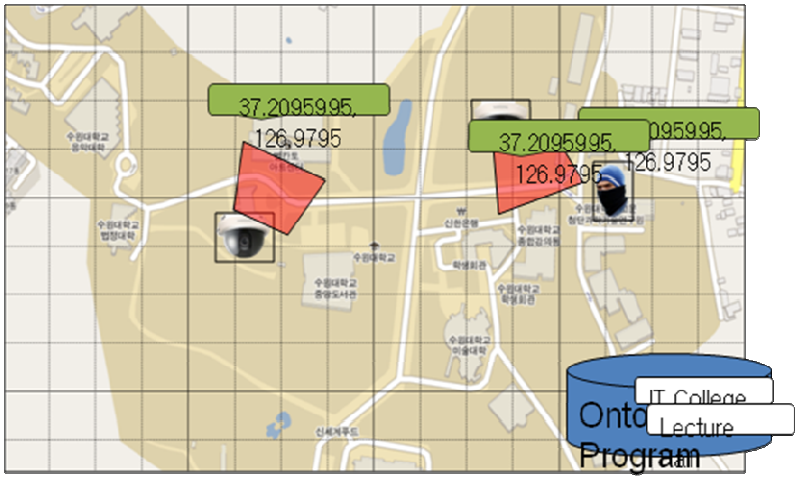

Figure 5. Subject tracking between two data source servers

6. CONCLUSION

We describe the profile data deploying guidelines for better context inference based on distributed ontology framework. Data source servers communicate each other freely within access control permission to perform their own intelligent distributed inference based on their own ontology knowledge base. Context inference including distributed multimedia data is widely used in distributed surveillance environment. In such a distributed surveillance environment, surveillance devices such as smart cameras may carry heterogeneous video data with different transmission ranges, latency, and formats. These devices not only can get services from a region server, but also they can form and generate a P2P network to provide services to each other. For such a P2P network, an effective profile framework that can handle heterogeneous devices is required. In this paper, we propose a flexible deploying scheme which is adaptive to the actual device demands and that of its neighbors. Our scheme uses conformity to update and share data in a cooperative way. Simulation studies are conducted to evaluate the effectiveness of our flexible profile data deploying scheme. Implementation is also going on into our distributed surveillance network environment. Based on mathematical derivations, the proactive deploying of profile data maximizes the reasoning engine's potential for making decisions from operating the hierarchical structure distribution. Profile data deploying structure, where data are deployed in the order of hierarchy and in the direction of suspect respectively, are formulated and optimal service policy is pursued in view of which middle agents to include in the system. Computational experiments are performed to investigate how the optimal deploying scheme and service policy responds to system parameter changes. The system parameters of our interest include data size, frequency, data update interval, the network distances between agents, and some relevant factors. Finally, simulation and implementation results are discussed in Section 5. More realistic data will be given in future work. Our scheme shows the efficiency of profile data deploying resulted in better context inference.

ACKNOWLEDGEMENTS

This work is supported by the GRRC program of Gyeonggi province. [GGA0801-45700, Center for U-city Security and Surveillance Technology]

REFERENCES

[1] A. Schlicht and H. Stuckenschmidt, (2008) "Towards Distributed Ontology Reasoning for the Web," WI-IAT08 pp. 536-539.

[2] J. Gomez-Romero, M. A. Patricio, J. Garcia, and J. M. Molina, (2011) "Communication in distributed tracking system: an ontology-based approach to improve cooperation," Expert Systems (The Journal of Knowledge Engineering), Vol. 28, No. 4, pp. 288-305

[3] L.Yin, G. Cao, (2006) "Supporting Cooperative Caching in Ad Hoc Networks," IEEE Tr. On mobile computing, Vol. 5, No. 1.

[4] Sem Borst, Varun Gupta, Anwar Walid. (2009) "Self-organizing algorithms for cache cooperation in content distribution networks," Bell Labs Technical Journal,

[5] Kam Fung Yeung, and Yanyan Yang. (2009) "Mobile information retrieval in a hybrid peer-to-peer environment, " Mobility '09 Proceedings of the 6th International Conference on Mobile Technology, Application & Systems

[6] K. Chitra and Dr. G. Padmavathi, (2011) "FAVQCHOKE: To Allocate Fair Buffer to a Dynamically Varying Traffic in an IP Network," International Journal of Distributed and Parallel Systems (IJDPS) Vol.2, No.1, pp.73-81

[7] U. H. Rao, (2011) "Challenges of Implementing Network Management Solution," International Journal of Distributed and Parallel Systems (IJDPS) Vol.2, No.5, pp.67-76

[8] ONVIF (Open Network Video Interface Forum) http://www.onvif.org

[9] PSIA (Physical Security Interoperability Alliance Specification Package Q12009, http://www.psiaalliance.org

[10] R. Martinez-Tomas, M. Rincon, M. Bachiller and J. Mira, On the correspondence between objects and events for the diagnosis of situations in visual surveillance tasks, Pattern Recognition Letters, vol. 29, Issue 8, pp. 1117-1135, (2008).

[11] Roberto Vezzani and Rita Cucchiara, (2010) "Video Surveillance On-line Repository (VISOR): an integrated framework," Multimedia Tools and Applications, Vol.50, No.2 pp.359-380

[12] T R Gopalakrishnan Nair, P Jayarekha, A Rank Based Replacement Policy for Multimedia Server Cache Using Zipf-Like Law, Journal of Computing, Volume 2, Issue 3, (2010)

[13] A. Paknikar, M. Kankanhalli and K. Ramakrishnan, (2000) "Caching and Streaming Framework for Multimedia," Proc. of the 8th ACM International Conference on Multimedia

[14] S. Yang, (2005) "An Efficient Access Control Model for Highly Distributed Computing Environment," Lecture Notes in Computer Science, Vol. 3741, pp.392-397

Authors

Soomi Yang received the B.S., M.S. and Ph.D. degrees in computer engineering from Seoul National University of Seoul, Korea, in 1985, 1987 and 1997 respectively. From 1988 to 2000, she was a researcher at Korea Telecom Research Center where she worked on telecommunication network, internet and information security. From 2000 to 2001, she was a visiting scholar at UCLA, USA. From 2002 to 2004, she was a faculty of the Suwon Science College. Since 2004, she has been on the Faculty of the University of Suwon, Korea, where she is a professor of computer science. Her research interests in information security include access control, network security, and secure system software.