

An Infrastructure for Efficient Reporting Workflow in Grid Based Teleradiology Architectures Using Relation Based Semantic Matching and Integer Linear Programming

Ayhan Ozan Yılmaz and Nazife Baykal
Department of Medical Informatics, Informatics Institute,
Middle East Technical University,
Ankara, Turkey
Email: e124572@metu.edu.tr, baykal@metu.edu.tr

Abstract—This paper proposes an infrastructure with a global workflow management algorithm in order to interconnect facilities and reporting units on a single access interface, decrease the access time of medical images and increase the efficiency of the reporting process. The inspection and radiologist attributes extracted by *Grid Agent* are modelled using a hierarchical ontology structure based on Digital Imaging and Communications in Medicine (DICOM) Conformance and DICOM Content Mapping Resource and World Health Organization (WHO) definitions. Attribute preferences rated by radiologists and technical experts or inferred by references are formed into reciprocal matrixes. Weights for entities are calculated utilizing Analytic Hierarchy Process (AHP). The assignment alternatives are processed by relation-based semantic matching (RBSM) and Integer Linear Programming (ILP). The results are evaluated based on turnaround time, workload and report quality and compared with the outcomes obtained by applying Round Robin, Shortest Queue and Random distribution policies.

I. INTRODUCTION

DUE to lack of radiologists within the facilities and consultation needs, the business model for radiology practice around the world is formed to include the facility's employing and outsourcing radiology services to non-local radiology groups [1]. Picture Archiving and Communication System (PACS) and Radiology Information System (RIS) are typically designed to handle local radiology communication and workflow management. However, remote accesses for non-local radiologists that serve several sites need to access medical images with a single interface and return medical reports efficiently, which requires speed, quality and workload optimization. Studies on PACS based on data grids [2], [3] propose co-allocation parallel transfer strategies to improve the non-local access interface and reduce the transfer time for medical images. Integration with heterogeneous resources and systems such as RIS and Hospital Information System (HIS) is also crucial for the quality of the service. This can be achieved by employing agents that support DICOM, Health Level 7 (HL7), Hypertext Transfer Protocol (HTTP), Cross Enterprise Document Sharing (XDS) and non-standardized data at regarding sites [4]. Turnaround time of a requested

report for an inspection is affected by the radiologist's availability, reporting speed and workload as well as the image transfer time. Therefore, workflow optimization should also be considered in the network and software architecture design. In previous research, multiple types of workflow optimization and semantic matching strategies are evaluated such as reinforcement learning [5], [6], machine learning (SVM, Bayes) [7] and relation based negotiation [8]. In this study, an infrastructure for medical image distribution is proposed and a RBSM algorithm enhanced by ILP is utilized to design medical image distribution strategy based on reporting workflow and efficiency. Subspecialty and quality of report are also critical parameters for teleradiology service. An inspection requiring subspecialty should be assigned to a radiologist with corresponding experience and high quality reports should be promoted in assignment process. In the proposed algorithm, experiences and subspecialties of radiologists are evaluated based on radiologist characteristics [9], [10] and report quality feedback [11] is included in the ontology map for the recalculation of weights by AHP.

II. METHODS

Workflow centric network architecture with an enhanced caching, querying and retrieving mechanism is implemented by seamlessly integrating *Grid Agent* and *Grid Manager* to conventional digital radiology systems. *Grid Agent* is deployed on each site which is responsible for rendering and transferring radiology data with PACS, RIS, and Workstations using DICOM protocol and with *Grid Manager*, clients using DICOM, HL7, HTTP, and Real Time Messaging Protocol (RTMP). *Grid Manager* is responsible for the flow management of images between sites and reporting units or distribution of reports based on the report distribution workflow algorithm. *Grid Manager* also enhances the image access time by providing non-local clients to query and retrieve medical images in parallel from multiple *Grid Agents* where medical content is cached during report distribution process. For web clients,

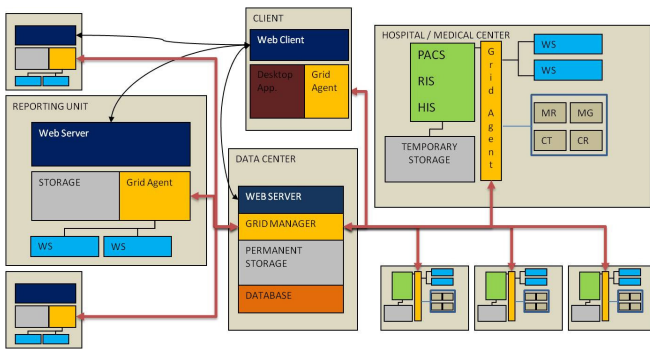


Fig. 1. Workflow centric network architecture with an enhanced caching, querying and retrieving mechanism

Java based open-source DICOM viewer software, *ImageJ*, is customized to stream image instances in parallel chunks.

A. The scenario

A typical data integration, communication and medical image delivery scenario starts with a non-local physician's request for a radiology inspection using the web interface. When the request is received, Grid Manager delivers the imaging request as an Extensible Markup Language (XML) message to the Grid Agent at the regarding medical center. Grid Agent informs the HIS and delivers the Modality Work List (MWL) request to the RIS. When the incoming patient is registered in HIS, Grid Agent is informed which afterwards gets the index of Grid Agents that have patient's data regarding previous examinations from the Grid Manager. Grid Agent pre-fetches the patient's previous medical information and synchronizes PACS and HIS in case a local radiologist examines the inspection. In parallel, Grid Manager automatically assigns the inspection to a remote radiologist after evaluating the experience, report quality, response time and technical adequacy parameters of registered radiologists and corresponding reporting units. Grid Agent at the reporting unit of the assignee receives the updated request list from the Grid Manager and fetches the patient's data including previous examination with the Grid Agent Index and synchronizes the data to PACS in the unit. The non-local radiologist can access the history of the patient independent from the vendor's software and the hospital where the data is acquired. The radiologist at the reporting unit retrieves medical images to be reported from several medical centers on a single interface and generates corresponding reports in RIS or using the web interface. The report is first delivered to the Grid Manager, then to the Grid Agent at the regarding medical center and finally to HIS.

B. The architecture

Teleradiology data serving and sharing architecture is composed of three main components: *Grid Manager* and *Data Center* forming the central node and *Grid Agent* forming the distributed network at the site nodes.

1) *The Grid Agent*: Grid Agent software is developed to run on an open source media server Red5 which also includes an embedded Tomcat servlet container for JEE Web Applications and supports streaming and shared object communication over RTMP. DICOM and HL7 messages are handled by asynchronous Java threads using dcm4che and HAPI Java libraries. The communication between Grid Agents and Grid Manager is accomplished using encrypted XML messages using HTTP and RTMP protocols.

2) *The Grid Manager*: Grid Manager is developed to run on Red5 and is specialized to send and receive encrypted XML and SOAP messages or DICOM files utilizing DICOM, HTTP or RTMP protocols. It communicates with Grid Agents and performs database, indexing and file operations at the center. Grid Manager has the Grid Index which includes the patient examination map cached by Grid Agents. The index is in shared object form so that a change in the index is pushed to all agents with the help of RTMP protocol. The caching mechanism at the agents provides the redundancy of the medical data so that the data achieve is distributed and web server maintenance costs are prevented.

3) *Data Center Architecture*: The central server is composed of the Application, Database and File Operation Layers. Grid Manager forms a bridge between these layers. The Database layer is implemented with open source Postgresql software. The database instances are implemented in shards to deliver large scale loads. Application layer is implemented with Red5 media server and File Operation layer is implemented with Tomcat Servlet Container.

4) *The clients*: Clients can query and retrieve medical images in parallel from multiple Grid Agents where medical content is cached during the pre-fetching and synchronization processes. For web clients, Java based open-source DICOM viewer software, *ImageJ*, is customized to stream image instances in parallel chunks. A query that is directed to the grid agent in a hospital by workstations is also directed to other grid proxies. Consequently, the query is performed at every hospital and central web server. Grid Manager provides the Grid Agent and consequently workstations with the result list and the images or data can be retrieved by the help of Grid Manager. Parallel downloading and efficient query algorithms in the Grid Manager enhances the bandwidth usage and time delay.

C. Medical Image Delivery Optimization

1) *Problem statement*: In order to claim that a radiologist is the optimum choice as a reporter for an inspection, parameters such as experience of the radiologist, response time, workload quota of the radiologist, technical adequacy of the reporting unit that the radiologist is located have to be evaluated.

Experience of the radiologist: Based on the expertise area of the radiologists or experiences on practice, radiologists may be better equipped in certain modalities, diseases or body systems. With reference to the studies that have been carried out on the association between radiologist characteristics and

interpretive performance of diagnostic radiology, a hierarchical structure is defined.

Response time: Response time is another important parameter that should be taken into account while estimating the most suitable reporter for the inspection. A radiology inspection for diagnostic purposed should be reported typically in 48 hours while an urgent inspection should be reported in at most 4 hours. The factors effecting the response time are inspection file delivery time depending on the inspection file size and reporting unit bandwidth, radiologist availability time based on the schedule and radiologist reporting time based on the modality, protocol and statistical data. The statistical data for response time is populated as the radiologists save their reports corresponding to inspections related with a certain set of modality, disease, body part and anatomy.

Workload quota of the radiologist: In order to achieve an efficient reporting process and to balance incomes, each radiologist should be assigned with inspections within certain workload limits. However, every reporting process is not equal in effort. The work load and payments of reporting processes are determined according to the "Performance Point Documentation (SUT)" announced by the Turkish Ministry of Health [13]. "Performance Point Extension Proposal" proposed by the Turkish Society of Radiology is used to strengthen the estimations on the average reporting time. In urgent cases, response time is much more important than the workload quota and expertise area; therefore, expertise area and workload quota are evaluated as secondary importance in emergency situations.

Technical adequacy of the reporting unit: Based on the inspection distribution scenario within this study, it is assumed that the radiologists are located in reporting units, where the assigned inspections are synchronized for access. Therefore, the technical infrastructure of the reporting unit effects the response time and the capacity of reporting service. Bandwidth of the unit affects the response time, while storage capacity and performance of the workstations determine the technical adequacy of the reporting unit. The medical monitor resolution is also taken as a requirement parameter as inspections of certain modalities need high resolutions for investigation.

2) **Rendering entities into ontology maps:** Ontology maps include the main nodes of Experience, Response Time, Workload and Technical as illustrated in Fig. 3. Experience is evaluated by the assessment of subquantities for each subnode Modality, Body Part, Anatomy and Disease. Similarly each node is connected hierarchically to subnodes having a weighted relation based on AHP. The input for the assessment process is provided by the inspection DICOM file. Each DICOM file provides entities that determine experience, response time, workload and technical requirements as illustrated in Fig. 2. *dcm4che* open-source DICOM Java library is used to render inspections in DICOM format. The modality of inspection, body part and anatomy examined, protocol requested, file size, series and slice numbers, resolution data are rendered into XML for RBSM and ILP processes. It is assumed that the pre-diagnosis is either embedded into the inspection or

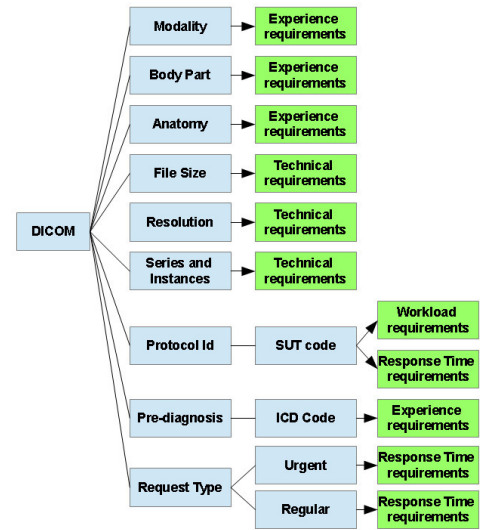


Fig. 2. DICOM file structure. The file structure is rendered to obtain components and these components are used to form the ontology map of the inspection DICOM file.

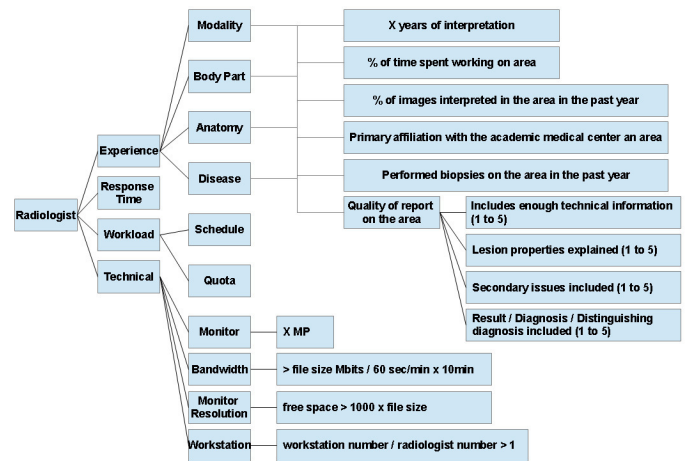


Fig. 3. Ontology map to calculate weights and ratings in inspection assignment to radiologists

entered manually by the report requester as 10th revision of International Classification of Diseases (ICD-10) code. The attributes rendered in the inspection files of DICOM format are used to derive the demand criterias for the radiologist.

3) **AHP Process:** The pairwise comparison for each entity of the problem is represented by a pairwise comparison matrix.

The weight w_l for each entity l at hierarchical level k of the ontology map including n entities is calculated using pair-wise

comparison matrix element a_{lm} by the following equation:

$$w_l = \frac{\sqrt[n]{\prod_{m=1}^n a_{lm}^{(k)}}}{\sum_{l=1}^n \sqrt[n]{\prod_{m=1}^n a_{lm}^{(k)}}}, \quad l, m = 1, 2, \dots, n \quad (1)$$

4) *RBSM Process*: Ratings are evaluated with a recursive relation based semantic matching process [12] by normalizing the sum of multiplicative weights with the following equation:

$$r_l^{(k-1)} = \frac{\sum_{m=1}^M w_m^{(k)} r_m^{(k)} q_m c_{lm}}{\sum_{l=1}^L \sum_{m=1}^M w_m^{(k)} r_m^{(k)} q_m c_{lm}}, \quad (2)$$

$l=1, 2, \dots, L$, $m=1, 2, \dots, M$ where L and M are the number of entities at levels $k-1$ and k of the ontology map respectively. q_m is equal to 1 if the qualitative or quantitative condition is satisfied by the potential assignee on entity m and 0 otherwise. c_{lm} is the binary value representing the presence of the connection between entities l and m in the ontology map.

5) *ILP Process*: The ILP process is formulated by defining the function to be maximized

$$J = \sum_{i=1}^S \sum_{j=1}^R r_{ij} x_{ij}, \quad i = 1, 2, \dots, S, \quad j = 1, 2, \dots, R \quad (3)$$

for S inspections and R radiologists where x_{ij} is 1 if inspection i is assigned to radiologist j or 0 otherwise; and r_{ij} represents the resultant rating for each potential assignment calculated by the *RBSM Process*. The constraints are defined as the following:

Assignment constraint: Each inspection is assigned to one and only one radiologist.

$$\sum_{j=1}^R x_{ij} = 1 \quad (4)$$

Workload constraint: The sum of assigned inspection estimated effort e_i should not exceed radiologist's workload l_j .

$$\sum_{i=1}^S e_i x_{ij} \leq l_j \quad (5)$$

Storage constraint: The storage free space s_j in the server at the reporting unit should be at least C_1 times that of the incoming inspection with file size f_i in bytes.

$$\sum_{j=1}^R \frac{s_j}{f_i} x_{ij} \leq C_1 \quad (6)$$

Bandwidth constraint: Transfer time of an inspection with file size f_i in bytes to a reporting unit with bandwidth b_j in bits per second should not take more than C_2 seconds.

$$\sum_{j=1}^R \frac{8 \times f_i}{b_j} x_{ij} \leq C_2 \quad (7)$$

Response time constraint: The response time which is equal to the sum of transfer time of the inspection file, radiologist

TABLE I
SIMULATED PRE-DIAGNOSIS ENTRIES AND MODALITY, BODY PART, ANATOMY INFORMATION RENDERED FROM THE DICOM FILES FOR 100 SAMPLE RADIOLOGY INSPECTIONS.

Ontology	Glossary
Modality	MR, CT, DX, MG
Body Part	Head, Leg, Throat, Knee, Shoulder
Anatomy	Skull, Leg, Upper Abdomen, Lower Abdomen, Shoulder, Spine
Disease	C39 : Malign neoplasm of inspiration system and inner thorax organs, C50 : Malign neoplasm of breast, C71 : Malign neoplasm of brain, C76 : Malign neoplasm of thorax, C78 : Malign neoplasm of rectum, D43: Neoplasm of brain and central neural system S02 : Skull and face bone fracture S42 : Shoulder and fore arm fracture S82 : Calf and knee fracture S83 : Dislocation, sprain or strain of knee and ligaments

availability time $t_{a,j}$ and radiologist's inspection specific average reporting time $t_{r,ij}$ should be less than the requested time $t_{req,i}$. This is typically 4 hours for urgent cases.

$$\sum_{j=1}^R \left(\frac{8 \times f_i}{b_j} + t_{a,j} + t_{r,ij} \right) x_{ij} \leq t_{req,i} \quad (8)$$

ILP process is implemented using the *lp_solve* Java library *OptimJ*.

D. Experimental Design

The performance of the proposed algorithm is tested using 100 sample radiology inspections. A simulation test bed is adopted with 4 imaging facilities and 3 reporting units, 1 data center and 2 non-local clients as virtual machines on different subnets. 6 radiologists working in 3 reporting units are registered and their experience, reporting unit technical capabilities are defined using the web interface. Round robin, random, shortest queue distribution policies are compared to RBSM and ILP distribution algorithms.

III. RESULTS

The results are evaluated based on experience rating, response time success rate and workload average deviation values.

Experience rating is a normalized value between 0 and 1 where higher experience rating is required for better reporting quality.

Response time success rate for policy p is defined as

$$sr_p = \frac{1}{S} \sum_{j=1}^R \sum_{i=1}^S s_{ij,p} x_{ij,p} \quad (9)$$

TABLE II

EXPERIENCE RATING, RESPONSE TIME SUCCESS RATE AND WORKLOAD AVERAGE DEVIATION VALUES FOR THE APPLIED DISTRIBUTION POLICIES: ROUND ROBIN, RANDOM, SHORTEST QUEUE, RBSM AND RBSM+ILP.

	Round Robin	Random	Shortest Queue	RBSM	RBSM+ILP
Experience Rating	0.34	0.37	0.40	0.75	0.71
Response Time Success Rate	0.74	0.72	1.00	0.93	0.99
Workload Avg. Dev.	2.73	3.60	1.94	0.49	0.28

$$\text{where } s_{ij,p} = \begin{cases} 1 & t_{ij,rep} \leq t_{i,req} \\ 0 & \text{otherwise} \end{cases}$$

for reporting time of inspection i by radiologist j , $t_{ij,rep}$ and requested reporting time for inspection i , $t_{i,req}$. The maximum possible value for sr_p is 1 and higher is the response time success rate, better is the distribution policy.

Workload average deviation for policy p , ld_p , is a measure of how efficient the radiologist resources are utilized indicating the distance from the load limit.

$$ld_p = \frac{1}{R} \sum_{j=1}^R \frac{\left| \sum_{i=1}^S e_i x_{ij,p} - l_j \right|}{l_j} \quad (10)$$

where e_i is the estimated time to report the assigned inspection i and l_j is the workload of radiologist j . The minimum possible value for ld_p is 0 which means the assignment workloads are equal to the defined workload limits for each radiologist. Therefore, distribution policy can be evaluated as more successful in terms of workload efficiency when ld_p approaches 0.

IV. DISCUSSION

The proposed architecture increases the efficiency of reporting process for teleradiology applications and provides a process centric network structure with an enhanced caching, querying and retrieving mechanism.

Shortest Queue policy has the highest response time performance; however it is inefficient in experience rating and workload distribution. Applying only RBSM gives the highest experience ratings, but integrating ILP with RBSM ratings provides a better response time success rate and the best performance for workload distribution with a small optimization trade off in experience rating. RBSM and ILP based image delivery also prevents bandwidth, storage or hardware related locks and latencies.

V. CONCLUSION

The proposed infrastructure decreases the storage costs, reporting costs, turnaround times and increases report quality and effectiveness of resultant treatments. The adaptation of medical sites and reporting groups to the architecture only requires the integration of Grid Agent into the present systems deployed on these sites which decreases integration costs and provides high interoperability.

The response time and report quality statistics for each radiologist are updated in real time. Therefore, it is considered that the proposed solution can be even more efficient and accurate in real case scenarios. Also the recalculation of weights based on the satisfaction level feedback for response time, report quality and workload distribution enhances the algorithm to make more accurate decisions.

REFERENCES

- [1] B.F. Branstetter, "Basics of Imaging Informatics: Part 2," *Radiology*, vol. 244, 2007, pp. 78-84, <http://dx.doi.org/10.1148/radiol.2441060995>.
- [2] C.T. Yang, C.H. Chen, M.F. Yang, "Implementation of a medical image file accessing system in co-allocation data grids," *Future Generation Computer Systems*, vol. 26, 2010, pp. 1127-1140, <http://dx.doi.org/10.1016/j.future.2010.05.013>.
- [3] C.T. Yang, M.F. Yang, W.C. Chiang, "Enhancement of anticipative recursively adjusting mechanism for redundant file transfer in data grids," *Journal of Network and Computer Applications*, vol. 32, 2009, pp. 834-845, <http://dx.doi.org/10.1109/ICPADS.2008.48>.
- [4] M. Benjamin, Y. Aradi, R. Shreiber, "From shared data to sharing workflow: Merging PACS and teleradiology," *European Journal of Radiology*, vol. 73, 2010, pp. 3-9, <http://dx.doi.org/10.1016/j.ejrad.2009.10.014>.
- [5] Z. Huang, X. Lu., H. Duan, "Mining association rules to support resource allocation in business process management," *Expert Systems with Applications*, vol. 38, 2011, pp. 9483-9490, <http://dx.doi.org/10.1016/j.eswa.2011.01.146>.
- [6] Z. Huang, W.M.P. Aalst, X. Lu, H. Duan, "Reinforcement learning based resource allocation in business process management," *Data and Knowledge Engineering*, vol. 70, 2011, pp. 127-145, <http://dx.doi.org/10.1016/j.datak.2010.09.002>.
- [7] Y. Liu, J. Wang, Y. Yang, J. Sun, "A semi-automatic approach for workflow staff assignment," *Computers in Industry*, vol. 59, 2008, pp. 463-476, <http://dx.doi.org/10.1016/j.compind.2007.12.002>.
- [8] W.K. Cheng, B.Y. Ooi, H.Y. Chan, "Resource federation in grid using automated intelligent agent negotiation," *Future Generation Computer Systems*, vol. 26, 2010, pp. 1116-1126, <http://dx.doi.org/10.1016/j.future.2010.05.012>.
- [9] D.L. Miglioretti, et al, "Radiologist Characteristics Associated With Interpretive Performance of Diagnostic Mammography," *Journal of National Cancer Inst.*, vol. (99)24, 2007, pp. 1854-1863, <http://dx.doi.org/10.1093/jnci/djm238>.
- [10] M. Eduard, et al, "Association between Radiologists' Experience and Accuracy in Interpreting Screening Mammograms," *BMC Health Services Res.*, vol. (8)91, 2008, pp. 1-10, <http://dx.doi.org/10.1186/1472-6963-8-91>.
- [11] J. Hohmann, et al, "Quality assessment of out sourced after-hours computed tomography reports in Central London University Hospital," *European Journal of Radiology*, vol. 81, 2012, pp. e875-e879, <http://dx.doi.org/10.1016/j.ejrad.2012.04.013>.
- [12] S. Colucci, et al, "A Formal Approach to Ontology-Based Semantic Match of Skill Descriptions," *Journal of Universal Computer Science*, vol. (9)12, 2003, pp. 1437-1454, <http://dx.doi.org/10.3217/jucs-009-12-1437>.
- [13] Performance Point Documentation (SUT), Turkish Ministry of Health, http://www.istanbul.saglik.gov.tr/w/sb/imis/pdf/ek_8.pdf.