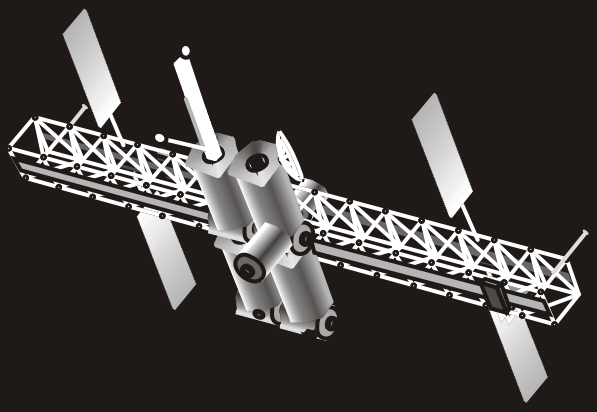
ENSC 803
Simon Fraser University
School of Engineering Science

2010 Course Proceedings



Form and Style: Graduate Publication in Professional Journals

Breast Cancer Scanning with Electrical Impedance Scanning

Daehan Chung,
Micro Instrumentation Laboratory, School of Engineering Science, 8888 university drive,
Simon Fraser University, Burnaby, Vancouver, BC, Canada, V5A 1S6

***Abstract*—Breast cancer is the most common cause of death in women. Although mammography technology improved survival rate, several limitations and problems still exist. In this paper, Electrical Impedance Scanning (EIS) system is proposed as an alternative way for detection of breast cancer. The previous limitations of existing EIS system are solved by applying polymer based microelectromechanical system (MEMS) technology. Especially, one of the MEMS technologies called soft lithography solved skin contact and low spatial resolution problems. This research aims to detect the size and location of breast cancer using improved EIS system.**

***Key Words*—Conductive nanoparticle composite polymer (C-NCP), electrical Impedance Scanning (EIS), impedance measurement, photo lithography.**

I. Introduction

A significant improvement in survival rate of women with breast cancer has been observed since mammography technology was developed. Nevertheless, breast cancer is still the most common cause of death in women between the ages of 45 and 55. Mammography exposes the breast to a small but significant amount of radiation, which can be a risky treatment, for young patients. This radiation can lead to breast cancer later stages of life. Furthermore, the procedure is generally considered to be uncomfortable and expensive.

In addition to mammography, palpation method or adjunct imaging methods such as ultrasound and magnetic resonance imaging are used to detect the breast cancer. However, high-quality mammography is still the most effective method presently available for breast cancer scanning.

Electrical Impedance Scanning (EIS) system is promising alternative way to detect the breast cancer. EIS system uses impedance analyzer machine which produces the impedance map of an object based on the spatial electrical characteristics throughout the volume of that object. The impedance analyzer uses very mild electrical pulses (AC current) to produce cole-cole plot (resistance vs. reactance graph) of the body being scanned. When the AC current passes through the body, the analyzer measures the impedance of the object, and sends the results to a computer. Then the computer divides the impedance into reactance and resistance, and draw cole-cole plot.

EIS is promising avenue of research to diagnose breast cancer less invasively compared to mammography because it does not pose any radiations or other known hazards. In addition, EIS enables diagnosis to be more accurate than mammography because cancerous breast tissue has higher electrical conductivity than normal tissue. Cancerous tissue will possess lower impedance than normal tissue impedance. If a patient has no cancer, the impedance of whole breast will present the same cole-cole plots. On the contrary, if the patient has breast cancer, the impedance of cancerous part will present different cole-cole plot. By comparing these cole-cole plots, we can detect the breast cancer.

Although EIS has great potential for breast cancer detection, it has not become a regular diagnostic tool due to some limitations such as low spatial resolution, poor electrode design, and skin contact issues. A large size, small number, and inflexibility of electrodes result in poor electrical contact with the skin and low spatial resolution.

We have developed conductive flexible electrodes using soft lithography technology to improve EIS system and solve the limitations mentioned above. Flexible electrodes arrays will be embedded in a bra-like garment and pressurized to conform to the breast. Once the electrodes are in contact with skin, impedance analyzer measures the impedance of tissue through the electrodes. Then a computer will change the results into cole-cole plots so we can compare the impedance of tissue easily. Fig. 1 indicates the overall system concept. The electrodes, impedance analyzer, and computer are connected to each other, and whole procedures are executed sequentially so it does not require long time.

The organization of this paper is as follows. In section II, the flexible electrodes are presented. This section demonstrates how to develop the electrodes, and why the electrodes are important in this project. In section III, proposed impedance measurements is presented to present how the system can be tested. Section IV provides expected results and discussion so we can check that the new system is applicable for real. Finally, in section V, we conclude this paper.

II. Flexible Electrodes

A. Making a molding master

A molding master is essential to fabricate electrodes because it has an electrode design on it. We can construct specifically designed electrodes using the molding master. As a condition of good molding master, it should be durable so that many electrodes can be made from it. Also, the electrodes' design on the mold should not be transformed or changed during electrodes making procedure. Several materials such as polymethylmethacrylate (PMMA) and SU-8 photoresist exhibit

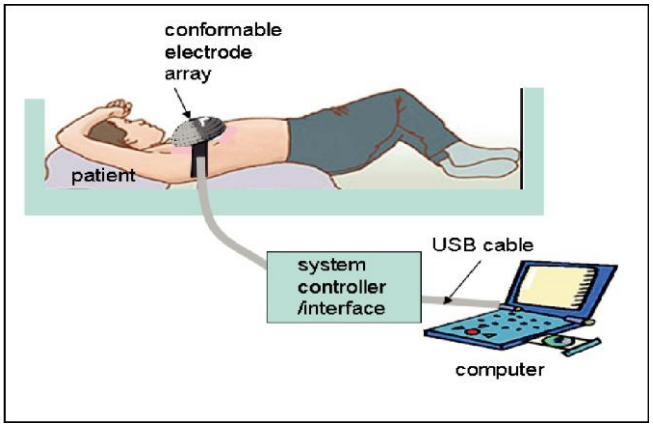


Fig. 1. Overall system concept.

Fig. 2. PMMA molding masters

good characteristics, but PMMA is cheaper and easier to handle compared to SU-8. In addition, PMMA mold is easier to manufacture than SU-8 mold.

We constructed a PMMA mold with a laser etching machine, which uses a laser to cut or etch substrate materials. The machine receives the image, which has to be etched, from the computer, and then etch the substrate using a laser with different power and speed. By having control on laser parameters such as the power and speed of laser, we can control the dimension of electrode design.

CorelDraw is a design tool for the laser etching machine. We designed 15mm × 5mm × 0.5mm paddle shaped electrodes using CorelDraw program. Fig. 2 indicates the picture of the mold. Head part will be attached to the tissue, and tail part will be connected to impedance analyzer through a wire. In actual application, the electrodes will only have size of several micrometers. The small size of electrodes will increase the spatial resolution and detect small cancerous tissues more accurately as a result.

B. Making an electrode

Polymer-based microelectromechanical system (MEMS) has advanced in recent years because of MEMS commercialization and influence from the flexible electronics industry. The most popular method for polymer microfabrication is soft lithography which involves heat curing of a flexible thermosetting polymer against a molding master. Using this technology, we developed flexible conductive nanocomposite polymers (C-NCPs) that are comprised of two polymer layers. One layer is flexible conductive polymer layer, and the other one is flexible insulated polymer layer.

Elastomers are chosen as a base polymer for flexible electrodes. An elastomer is a polymer with the property of viscoelasticity, which means they can deform reversibly under stress. The extreme flexibility of the elastomer enables the electrodes to bend freely. It will also increase the resolution and surface contact density, and thereby increase the lower detectable limits. However, elastomer is not electrically conductive.

Nonconductive polymer can be converted to conductive polymer by dispersing suitable nanoparticles such as carbon nanotubes or metal nanoparticles. We chose silver as a conductive nanoparticle because silver is cheaper than gold or platinum, and the most common metal nanoparticles that has a very low electrical resistivity, 1.617×10^{-8} Ω-m at 298°K. High conductivity of electrodes enables the current to flow well through the electrodes.

Fig. 3 indicates the procedure of making C-NCP electrodes. After making a molding master, we poured the mixture of silver nanoparticle and elastomer on the master, and scraped out unnecessary part of the mixture. Once the mixture has been molded, a nonconductive polymer was poured and cured on it to serve as a substrate. We used PDMS as a substrate polymer because it is also elastomer based polymer. In addition, PDMS is biocompatible, and adhere to the C-NCPs well. The entire assembly then was peeled off from the molding master.

C. Uniform Conductivity of Electrodes

The electrode should possess stable and uniform conductivity. The impedance, which is measured by impedance analyzer, is sum of impedances from tissue, electrodes, and wires which connect the machine and electrodes. If the conductivity of electrodes is unstable and varies, the impedance analyzer can't measure tissue impedance properly.

Acquiring uniform conductivity depends on dispersion of silver nanoparticle in the elastomer. The more nanoparticles disperse evenly, the more electrodes acquire uniform conductivity. A polymer consists of monomers, which are linked together firmly to form the polymer. As a result, dispersing conductive nanoparticles into the polymer is difficult because nanoparticles don't have enough space to disperse into and bond to the monomers. The nanoparticles must be very well dispersed, and ultrasonic mixing is used.

We stirred the mixture using both ultrasonic stirrer and magnetic stirrer. Using an ultrasonic stirrer is the most popular technique because it is quick, easy to use, and the dispersion is good as compared to other agitation methods. However, sometimes the polymer or nanoparticles settled to the bottom while ultrasonic stirrer was working. To prevent this precipitation, we used magnetic stirrer. A magnetic bar on the bottom rotates continuously, and prevents the polymer or particles from precipitating. After ultrasonic and magnetic

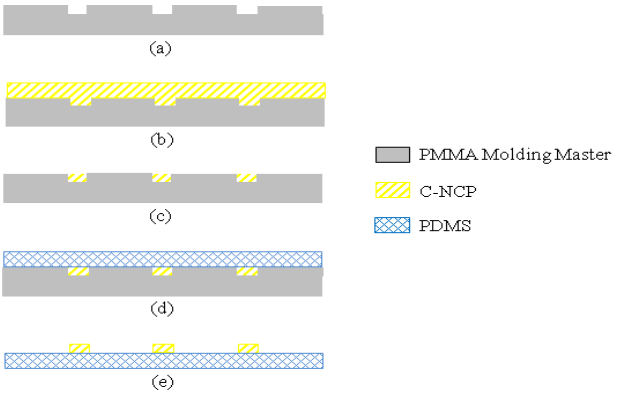


Fig. 3. Cross-sectional schematic diagrams of the C-NCP electrodes making process. (a) Making a molding master. (b) Pouring C-NCP mixture on the mold. (c) Scrape out unnecessary part. (d) Pouring PDMS on the mold. (e) Peel off the assembly from the mold.

agitation, we stirred the mixture manually for two minutes again, and poured on the PMMA mold master to develop conductive layer of electrodes.

D. Percolation

The electrical conductivity of C-NCPs is strongly dependent on the nanoparticles' density. At low nanoparticles density, C-NCPs behave like insulated polymer with very low conductivity. On the contrary, very high density of nanoparticles has detrimental effect on polymer's mechanical properties. The polymer will have lower flexibility, or will not cure because of high density of nanoparticles.

To present the nanoparticle density, weight percent is used. Weight percent is the weight of solute divided by the weight of solution multiplied by 100%. In N-CNPs' case, weight percent indicates the fractional amounts of nanoparticle out of C-NCPs' total weight. Percolation threshold is the point at which the first continuous chain of linked conductive particles is present. The weight percentage of silver nanoparticles should be larger than the percolation threshold to fabricate C-NCPs.

We could construct C-NCPs which have resistivity about 2×10^{-1} Ω-cm at 57 wt-% of silver nanoparticle. However, the resistivity of each electrode was not same. The electrodes had resistivity difference of ± 8% range. This variation is caused because the electrodes thickness was slightly different. The thickness shrank when heptane evaporated, and each electrode indicated different thickness from 0.45mm to 0.49mm.

E. Chloridation of electrodes

Water is the most common molecule in the human body, and hence it is the most abundant solvent. Most chemical reactions involve molecules dissolved in water, which is called ion

electrolyte. Not only chemical reactions, but also current in the body is carried by ions. Thus, the electrodes should serve as a transducer because current in the electrodes is carried by electrons. The electrodes should be able to change an ionic current into an electronic current, and vice versa.

Generally, two types of electrodes are possible: perfectly polarizable electrodes, and perfectly nonpolarizable electrodes. Difference between two types of electrodes depends upon what

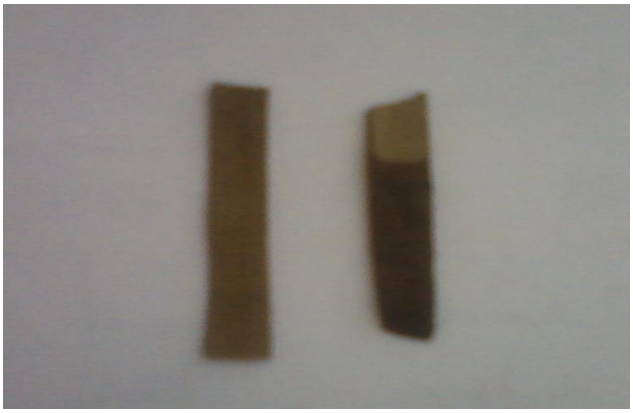


Fig. 4. A piece of C-NCPs before Chloridation (Right), and chloridized electrode (Left).

happens to the electrodes when a current passes between them and the electrolyte. Perfectly polarizable electrodes change their potential when current is applied, and act like capacitors. Hence, current can't pass the electrode-electrolyte interface freely and flow upon a change of potential. Normal metal electrodes are one example of perfectly polarizable electrodes so they are not suitable for tissue impedance measurement. On the other hand, the potential of the perfectly nonpolarizable electrodes will not change from its equilibrium potential even when a large current is applied. That is, current can pass freely across the electrode-electrolyte interface without any energy change to occur the transition. Therefore, perfectly nonpolarizable electrodes are the suitable electrodes for the EIS system.

A silver/silver chloride (Ag/AgCl) electrode is a practical electrode that satisfies the characteristics of a perfectly nonpolarizable electrode and easy to fabricate in the laboratory. Fig. 4 presents the picture of electrode before and after chloridation. The color of the electrode became darker after chloridation, but the conductivity of electrode hardly changed.

Chloridation also prevents the oxidation of silver in the air. The silver particle in the electrodes oxidized when they are exposed to the air. Once the silver particles are oxidized, the electrodes become nonconductive. Silver chloride (AgCl) compound prevents silver particles from oxidation so the electrodes can maintain their conductivity.

III. Measuring Impedance

(The result is not from C-NCP electrodes, but from wire electrodes. We couldn't test C-NCP electrodes because the impedance analyzer is out of function now)

To measure the impedance of tissue, we developed a fake tissue called phantom which is constructed from a mixture of gelatin and salt water. 6 electrodes were used to measure the impedance of phantom. The electrodes were attached to the center and periphery of phantom at the same angle and connected to the HP 4192A impedance analyzer. The machine measured the impedance of phantom between the central ground electrode and each of the remaining electrodes at a given frequency and sent the data to the computer. Then, the computer collected the data and converted it into cole-cole plot. We performed frequency sweeps over the frequency range of 10 KHz to 13 MHz in 10 KHz steps to simulate baseline tissue

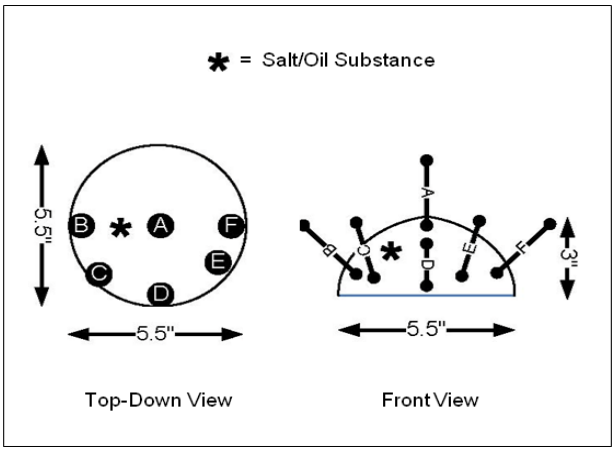


Fig. 5. Electrodes layout to detect anomaly in a breast phantom.

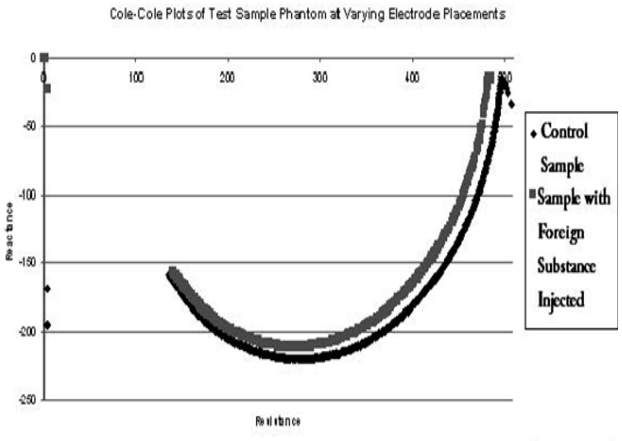


Fig. 6. Cole-Cole plot generated from breast phantom.

measurements. After measurements were completed, an artificial malignancy consisting of olive oil and salt was injected into the phantom, and the same measurements were repeated. Fig 5 indicates the location of electrodes on the phantom, and anomaly.

IV. Results

Fig. 6 demonstrates the cole-cole plot of phantom after impedance measurements. X-axis presents resistance, and Y-axis presents reactance. In the Fig. 6, the low frequency measurements took place near the upper-right portion, while the higher frequency measurements were towards the left.

The upper line indicates the impedance between electrode A and B. An anomaly was located between them, and the result presents that only the impedance between A and B differed from others, while the impedances between A - C, A - D, A - E, and A- F (lower line) were all the same. From the cole-cole plot, we could see that the anomaly was between electrode A and B.

In addition, we could check that the biggest impedance difference between the anomaly and normal phantom took place at relatively low frequency region (upper-left region). This phenomenon means we do not have to use large range of frequency sweep. The frequency range from 5 Hz to 6 KHz will be enough for detection of the breast cancer, and this narrow frequency range will further reduce data acquisition times. Narrow frequency sweep will have less effect on human tissue than large frequency sweep as well.

V. Conclusion

In this paper, we measured the impedance of breast phantom using the EIS system to detect the breast cancer. Our method could improve the EIS system and solves several problems that challenged the system in past by developing conductive nanoparticle composite polymers (C-NCPs) with soft lithography technology. We could expect that the C-NCPs will solve skin contact problem and increase spatial resolution. This increased spatial resolution enables doctors to detect very small cancerous tissues. In addition, the proposed new system requires less time than mammography, and obviously less invasive because it doesn't use any radiation.

We could check that the new system can work properly by measuring impedance of fake tissue phantom with several electrodes. The result proved that the anomaly had lower impedance than normal phantom. We could also find the location of the anomaly.

In actual application, we will use several hundreds of very tiny electrodes array, and the measured impedances will be changed into three dimensional images (3D image) so we can check the location and size of breast cancer easily.

Acknowledgment

I thank Bonnie L. Gray, and Ajit Khosla for insightful discussions. Help and advice of Ram Ramaseshan is gratefully appreciated. Parts of this work have been supported by NSERC, and Fraser Valley Cancer Center.

References

[1] John G. Webster, "Medical instrumentation: application and design" 4th ed, Hoboken, NJ: John Wiley & Sons, 2009, Ch.4.

[2] A. Khosla, B. L. Gray, "Preparation, characterization and micromolding of multi-walled carbon nanotubes polydimethylsiloxane conducting nanocomposite polymer", Materials Letters. 2009. 63: 1203-1206.

[3] Younan Xia, George M. Whitesides, "Soft Lithography", Annu. Rev. Mater. Sci. 1998. 28: 84-153.

[4] Matthew Giassa, A. Khosla, B. L. Gray, "Application for low frequency impedance analysis systems", Journal of Electronic testing, 2010. 26: 139-144.

[5] K. S. Kohli, D. V. Rai, V. K. Jindal, and N. Goyal, "Impedance of goat eye lens at different DC voltages", medical and Biological Engineering and Computing, vol. 36, pp. 604-607, Sep. 1998.

[6] B. L. Gray, A. Khosla, "Microfabrication and applications of nanoparticle doped conductive polymers", unpublished

[7] Nathalie K. guimard, Natalia Gomez, Christine E. Schmidt, "Conducting polymers in biomedical engineering", Prog. Polym. Sci. 2007. 32:876-921.

Daehan Chung received his bachelor's degree in Electrical engineering from Korea University in 2008, and is currently a M.A.Sc candidate in School of Engineering Science in Simon Fraser University. His research interests include flexible electrodes for impedance analysis of cancerous tissue, and nanofabrication.

Analyzing the correlation between High-Level Information and FPGA Placement

Farnaz Gharibian
School of Engineering Science
Simon Fraser University

***Abstract*—One of the most popular placement algorithms for Field Programmable Gate Arrays (FPGAs) is called Simulated Annealing (SA). This algorithm creates a good quality placement from a flattened design that no longer contains any high-level information related to the original design hierarchy. Unfortunately, placement is an NP-hard problem and as the size and complexity of designs implemented on FPGAs increases, SA does not scale well to find good solutions in a timely fashion. As modern FPGAs can be used to implement Systems- and Networks-on-Chip, designers are required to spend an increasing amount of time waiting for place and route tools to complete. The waiting time does not matched by an increase in the power of computing work stations.**

In this paper, we investigate if high-level information can be reconstructed from a flattened netlist and evaluate how that information is realized in terms of its locality in the final placement. If there is a strong relationship between good quality placements and high-level information, then it may be possible to divide a large design into smaller components and improve the time needed to create a good quality placement. Our preliminary results suggest that the locality property of the information embedded in the high-level Hardware Description Language (HDL) structure (i.e., "module", "always", and "if" statements) is greatly affected by both the designer and the design itself. A reconstructive algorithm, called affinity propagation, is also considered as a possible method of generating a meaningful coarse grain picture of the design.

I. Introduction

Field Programmable Gate Arrays (FPGAs) have gained popularity as an implementation platform as their logic density has grown, allowing them to implement complete Systems-on-Chip (SoCs). During the development of an automated Computer Aided Design (CAD) flow for FPGAs, it was recognized that the placement problem could only be solved in Non-Polynomial (NP) time [1]. To this end, CAD designers leveraged heuristics to obtain good quality placement solutions in a timely fashion.

One of the more popular heuristic placement algorithm implementations for FPGAs is called Simulated Annealing (SA). It tries to create a good quality placement from a flattened design that no longer contains any high-level information from the original design hierarchy. As the size and complexity of FPGA designs increases, SA does not scale well; in fact, the time designers are required to wait for the successful place and route of a design using commercial CAD tools is reportedly not being alleviated by the gained computing power in new work stations [2].

In this paper, we present a study to determine the quality and types of high-level information that may be reconstructed by SA in the final placement; to quantify the quality of this reconstruction, we use Manhattan Distance (MD) as a measure of locality. The specific contributions are:

- An analysis that suggests that the quality of the information embedded in the high-level circuit structure (i.e., *module*, *always*, and *if* statements) is greatly affected by both the designer and the design itself.
- How clustering algorithms, such as affinity propagation [3], can reconstruct high-level information that may be used to create coarse grain components to provide the flexibility needed for multi-phased placement.
- A method of evaluating the quality of high-level information in relation to placement using MD and the distribution of high-level granular blocks. This methodology allows new reconstructive algorithms to be easily analyzed and compared to one another.

If some general relationship between good quality placements and high-level information exists, then dividing a large design into smaller components may improve the speed with which a good quality placement is obtained as we could better leverage new multicore processing platforms. Ideally, the general placement of a design using SA can be somewhat intuited from the cost function it uses to evaluate the quality of a placement: minimize the wire lengths of connected blocks with the objective of minimizing the critical path. However, as SA relies on randomly swapping clusters, there is no direct link between the original design structure and the structure of the final placement.

The remainder of this paper is organized as follows. Section II discusses the related work including placement algorithms for FPGAs and using some form of design partitioning/clustering in combination with placement. Section III describes what we mean by "high-level information" and why partitioning designs into coarse grain components guided by high-level information may be valuable for reducing run time and even improving placement quality. Section IV explains the tools and algorithm used in this paper and Section V summarizes the CAD flow used to obtain our data. Section VI presents our experimental results and Section VII concludes the paper, with thoughts as to possible directions for future work.

II. Background

The goal of placement in the FPGA CAD flow is to find a valid placement for each logic block while trying to minimize the critical path; this objective is typically represented as a cost function used to minimize the total length of interconnect required. There are two main approaches to placement: analytical and heuristic algorithms. Analytic placers, such as quadratic placement [4] [5], are commonly used in commercial FPGA CAD flows to provide an initial high-level placement. They generally complete in less time than other heuristics but do not guarantee a legal placement as the final placement might have overlapping logic blocks [4]. Instead, they are generally used to provide an initial placement for larger designs that is fine tuned using heuristic algorithms (e.g., SA) to remove the overlap, resulting in a more scalable run time for placers as design complexity increases.

Simulated Annealing [6] (SA) is often used for at least a portion of an FPGA's placement. Unfortunately, SA requires significant computation time to produce good placements. SA swaps random clusters to move through the design space, in combination with a "cooling table" that allows the algorithm to select the occasional "poor" placement (in terms of its cost function) to perform some hill climbing. Different techniques have been used to improve SA's run time, some of which have been included in the version of SA incorporated into the academic CAD flow VPR [7]. Other techniques for reducing the run time of SA rely on:

- some form of partitioning/clustering the design into groups of clusters (which we call ***super-clusters***) before running the placer [8],
- quickly creating a *good* initial placement to reduce the iterations required to find the final placement [9], and
- using multi-threading techniques that allow the placer to leverage the increased processing power available on multicore processors [10], [2].

Although our long term research objectives include creating a placement algorithm that scales better in performance for multicore processors, the current study focuses on a *good* placement's final structure in relation to the actual design. As such, the remainder of this discussion on SA focuses on previous work reducing the actual number of iterations of SA using partitioning/clustering or a good initial placement. Work on multi-threading techniques is complementary to this work as it instead focuses on completing SA iterations more quickly.

FPGA placements based on SA use random initial placements and do not consider information embedded in the original design as circuits are typically flattened before SA is run. Some previous work in ASIC placers suggest that high-level information and design hierarchy should be considered during both clustering [11] and placement [12]. Previously, floorplanning (or hierarchical) approaches to placement, based on the design's hierarchy as specified in its RTL have been introduced [13], [14]. For example, Emmert et al. [13] start by decomposing the FPGA device into an array of placement bins with the same physical dimensions. Then a macro-based netlist of soft and hard macros targeted to the device is created. Finally, it groups the macros into clusters that can be mapped into bins.

III. High-Level Information

The idea of trying to replace SA's time consuming step of moving from a random placement to a reasonable one via random swaps with a good initial placement is discussed by Grover [15]. This Two Stage Simulated Annealing (TSSA) work is based on the hypothesis that: SA wastes effort using random swaps to generate its first reasonable solution; instead, it would be better to generate a good initial placement utilizing more computationally efficient methods and apply SA at a low temperature. We propose using *high-level Information* to guide the generation of this initial placement, where we define *high-level Information* as: a granular grouping of the design based on connectivity and/or structure that may be obtained either from its graphical representation or its original HDL (design-level) description. We are particularly interested in determining if this information can be used to create coarse grain substructures to guide initial placements for heuristics, and potentially lead to a more scalable placement solution multicore architectures.

To evaluate how useful this approach might be, we first assume that as SA generates good quality placements, it uncovers substructures that have locality. We propose that these "uncovered" substructures fall into two categories: 1) those that are tied to the original design's structure and can be found using high-level information, and 2) those that are not apparent from the original design but truly "uncovered" by SA's random swaps. The remainder of this paper focuses on the first category of substructures, investigating what types of high-level information SA "reconstructs" in its final placement and the quality of reconstruction. To quantify this metric, we utilize the Manhattan Distance (MD) of the set of clusters that implement the substructure, where a smaller MD reflects greater locality for the substructure and thus a better reconstruction. If certain types of information have poor locality in an SA placement, it may not be as useful in creating coarse grain substructures for an initial placement.

IV. Experimental Approach

In this study, we investigate the locality of high-level information within a final SA placement using two approaches. The first is based on the structure of the source HDL, where we try to determine how hierarchical HDL structures (specifically, *module*, *always*, and *if* blocks) are reflected in the final placement. An academic syntheis tool (ODIN-II) used to detect and extract HDL components is reported in Section IV-A. The second approach uses a reconstructive algorithm, called affinity propagation [3], to analyze the connectivity graph of the netlist after logic has been packed into clusters. The algorithm is described in Section IV-B which detects relationships that determine what clusters should be considered part of the same coarse grain substructure.

```
module test_simple_fsm (Clock, Resetn, w, z);
   input Clock, Resetn, w;
   output z;
   reg [1:0] y, Y;
   parameter [1:0] A = 2'b00, B = 2'b01, C = 2'b10;

   always @(w or y)
   begin
      case (y)
         A: if (w)
               Y = B;
            else   Y = A;
         B: if (w)
               Y = C;
            else   Y = A;
         C: if (w)
               Y = C;
            else   Y = A;
         default:   Y = 2'b00;
      endcase
   end

   always@(posedge Clock)
   begin
      if (Resetn == 1'b0)
         y <= A;
      else y <= Y;
    end

   assign z = (y == C);
endmodule
```

Fig. 1. An example of a verilog file

A. ODIN-II

Odin [16] is an open source academic synthesis tool for heterogeneous FPGAs and permits flexible targeting of hard structures in FPGAs. Different steps are taken to map verilog design to two different FPGA CAD flows: Altera Quartus II and the academic VPR CAD flow. First, a parser parses the input file to a structure. Second, a flat netlist is created that consists of structures including logic blocks, arithmetic operations, registers. Third, synthesis and optimization is performed on the netlist. Fourth, an inferencing stage searches for the structures that could be mapped on hard circuits on the FPGA. Finally, nodes in the netlist are mapped to available resources on the FPGA. In this paper, we have used ODIN-II which is the successor of Odin. The original Odin used Icarus [17] as the front-end parser while in ODIN-II [18] an Abstract Syntax Tree (AST) structure is created from a verilog input file. AST captures the fundamental structure of the input in a tree form, while omitting unnecessary syntactic details [19]. Three major tree nodes that are used in this paper are discussed in following sections.

1) MODULE: The verilog file may contain different modules. Number of modules is kept in a global variable named *num_modules*. Figure 1 shows an example of an input file to ODIN-II. Tree structure created for this file is kept in ast_modules[0] which is the root of the tree. The root contains three children: children[0] contains the information about the module (e.g., module name), children[1] contains information about module lists (lists of input and output signals to the module), and children[2] contains information about module items. For example, the name of the module ("test_simple_fsm") is accessible through $ast\_modules[0]- > children[0]- > types.identifier$. Each child of the tree is also a subtree itself. In each node, number of children is available in *num_children* variable. In the example shown in Figure 1, number of module items is available in $ast\_modules[0]- > children[2]- > num\_children$, which is 7. The first module item is an *input* declaration containing 3 children ($ast\_modules[0]- > children[2]- > children[0]- > num\_childen$).

```
if (w)
   Y = B;
else   Y = A;
```

Fig. 2. An example of a IF statement in verilog

2) ALWAYS: Each *always* node has two children: children[0] contains the information about sensitivity list parameters, and children[1] contains statements of always block. For example, children[0] of the *always* node of the following example has 2 children: "always @ (w or y)". The example shown in Figure 1 has two *always* nodes which the first one can be accessed using $ast\_modules[0]- > children[2]- > children[4]$.

3) IF: Figure 2 shows an if statement in verilog. The tree node created for each "if" statement has three children: children[0] contains the if condition (in this example checks the value of *w*), children[1] contians the statement for the case that the condition is true, and children[2] contains the statement for the case that the condition is false. For example, false condition (Y=A) node is accessible through $if\_node- > children[2]- > children[0]$.

B. AFFINITY

Affinity propagation is a clustering algorithm that is applied to an input graph to group clusters based on a defined similarity factor and generate a set of super-clusters. These super-clusters are created via a series of message passing steps, where nodes in the graph communicate their attraction to other nodes. The number and size of the super-clusters created by the algorithm depends on the structure of the graph and a similarity factor that describes how every two nodes in the graph are initially related to one another. We have implemented two versions of affinity propagation that use slightly different metrics to create the initial similarity factors. The first (AFFIN1) assumes that the similarity of two nodes in the graph (representing circuit components) is higher for shorter graph distances.

For example, if two nodes are connected directly to each other, the assumption is that they should be placed closer together on the FPGA. However, for longer path lengths, the similarity between nodes is lower as they do not need to be placed close together. Our second version of affinity (AFFIN2) creates similarities based on three properties: common connections, proximity, and fan out. Specifically, a node is considered more similar to another node by this version of the

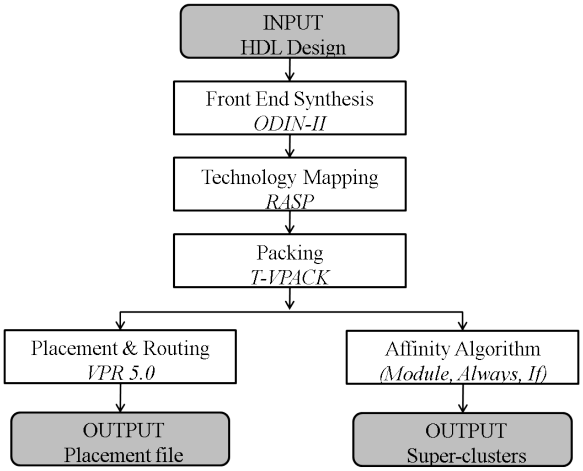


Fig. 3. FPGA CAD flow used in this research

TABLE I
THE FPGA ARCHITECTURAL PARAMETERS

Parameter	W	N	K	F_{cin}	F_{cout}	F_s	routing
Value	20% larger than minimum	10	4	0.18	0.1	3	uni-directional

algorithm if: they have more than one connection, are close in proximity, and their respective fanout is high. For example, unlike AFFIN1, AFFIN2 is more likely to closely place two nodes with many connections than two nodes with closer proximity yet having only one connection between them.

V. EXPERIMENTAL METHODOLOGY

Figure 3 shows the FPGA CAD flow used in this work to determine how structural information from both the HDL design and the circuit netlist are related to the SA's final placement. A Verilog HDL design is inputted to Odin-II [18], which creates a flattened netlist consisting of system-level I/Os, logic gates, flip-flops, and hard circuits (e.g., multipliers/memories). This netlist is passed into RASP [20] for logic optimization and mapping to Look-Up Tables (LUTs). RASP was chosen specifically for this project as it maintains the majority of design-level naming in the netlist, allowing us to measure the MD of various structures (i.e., *module*, *always*, and *if*) in the final placement. Next, T-VPack [21] takes the output from RASP and packs the LUTs and registers into clusters. The output of T-VPack is then placed and routed onto an FPGA architecture described by the parameters summarized in Table I using VPR 5.0 [22]. Our two versions of affinity propagation (AFFIN1 and AFFIN2) also use T-VPACK's output to generate super-clusters based on their respective similarity factors outlined in Section IV-B.

We use the Manhattan Distance (MD) of a super-cluster's bounding box in the final SA placement to measure the locality of a super-cluster. Specifically, a super-cluster's MD is equal to half the perimeter of the bounding box of all the clusters in that super-cluster. The clusters in the super-clusters (generated using affinity propagation or from a benchmark's *module*, *always*, and *if* blocks) are located basically the same way

TABLE II
RESULTS AFTER VPR COMPLETES PLACE & ROUTE

Benchmark	Number of IO	Number of Clusters	Grid Size	Std Dev	GMEAN
cfc18	111	627	26*26	5.08	95.66
cft8	69	1031	33*33	5.59	81.58
desa	190	265	17*17	4.39	111.59
iir1	59	142	12*12	1.02	54.75
oc54	140	530	24*24	0.64	24.60
pajf	103	166	13*13	14.44	233.50
rsd2	32	496	23*23	0.88	57.61
dconvert	258	445	22*22	0.74	20.28
dsystemC	162	409	21*21	0.39	17.98
glue2	40	43	7*7	3.13	91.75
rsd1	20	192	14*14	1.23	131.14

- by searching the final placement for the locations of all clusters identified in a super-cluster. However, whereas the affinity propagation algorithm outputs a list of super-clusters and their component clusters, we use the design-level naming scheme generated by Odin-II and passed through the CAD flow to locate individual *module/always/if* instantiations.

VI. RESULTS

In this section, we summarize the outputs of the VPR place and route and discuss how certain types of high-level information correlate to the final placement.

A. VPR results

Table II shows the VPR statistics for 11 of the larger open source benchmarks available in Verilog HDL. Column 1 contains the names of the benchmarks, and Columns 2 through 4 list the number of I/Os, the number of clusters, and the array size for each benchmark respectively. Columns 5 and 6 show the standard deviation and geometric mean of the operating frequency of the benchmarks in MHz over ten random seeds.

B. Affinity and HDL Placement Correlation Results

Table III summarizes the results of our first experiment using AFFIN1 (top) and AFFIN2 (bottom) to reconstruct high-level information into super-clusters. First, we run both AFFIN1 and AFFIN2 to create the super-clusters and then calculate the MD of these super-clusters in the ten placements generated for each benchmark to evaluate our affinity propagation clustering algorithms. Columns 2 shows the number of super-clusters constructed by each version of the algorithm. Columns 3 through 6 gives the distribution of clusters in the super-clusters where the 1st two columns list the minimum and maximum number of clusters in a super-cluster and Columns 5 and 6 show the standard deviation and the geometric mean. Columns 7 through 10 list the distribution of super-cluster MDs (Local MDs) in a similar format. Columns 11 and 12 give the standard deviation and geometric mean of the Global MD generated by totalling the MDs of each super-cluster. Column 13 contains the ratio of the average Local MDs for the super-clusters (Column 10) divided by the average number of clusters per super-cluster (Column 6).

Table IV shows the results of our second experiment, measuring how the *module*, *always*, and *if* HDL constructs

TABLE III
RESULTING SUPER-CLUSTERS USING TWO DIFFERENT AFFINITY PROPAGATION METRICS.

AFFIN1												
Benchmark	Super-clusters	Clusters				Local MD				Global MD		Local MD/Clusters
		min	max	STD	Gmean	min	max	STD	Gmean	STD	Mean	
cfc18	65	1	179	21.52	6.71	0	49	8.6	16.6	23.44	1041.8	2.47
cft8	91	3	516	53.51	5.84	2	63	12.07	14.35	25.34	1332.6	2.46
desa	6	6	231	91.53	12.22	5	31	10.61	9.33	1.78	57.5	0.76
iir1	15	5	51	11.57	7.32	6	22	4.99	11.2	6.28	169.1	1.53
oc54	43	3	247	36.72	6.88	2	45	7.88	17.14	12.15	737.4	2.49
pajf	5	3	148	64.18	8.82	3	24	7.92	10.8	2.66	54.2	1.22
rsd2	50	3	209	28.78	6.03	3	43	7.51	10.94	14.33	563.3	1.81
dconvert	38	3	228	36.08	6.16	3	41	8.35	13.76	12.64	520.5	2.23
dsystemC	37	3	222	35.68	5.48	2	39	8.16	11.73	14.61	414.7	2.14
glue2	6	3	21	6.94	5.46	3	11	3.14	5.67	3.38	34.1	1.04
rsd1	19	3	88	19.01	6.2	2	26	6.18	10	9.69	190.8	1.61
AFFIN2												
Benchmark	Super-clusters	Clusters				Local MD				Global MD		Local MD/Clusters
		min	max	STD	Gmean	min	max	STD	Gmean	STD	Mean	
cfc18	13	18	75	17.67	44.8	13	46	8.35	32.23	9.52	435.8	0.72
cft8	35	12	131	21.78	25.06	9	62	15.22	24.51	21.26	860.9	0.98
desa	8	17	64	19.45	29	11	22	3.72	17.88	6.66	139.9	0.62
iir1	8	7	31	7.7	16.17	5	21	5.28	15.25	5.82	113.5	0.94
oc54	18	9	52	12.89	26.46	13	31	5.04	23.17	6.11	417.4	0.88
pajf	118	1	40	3.63	1.07	0	15	1.77	0.3	2.67	38	0.28
rsd2	16	5	204	52.1	13.99	4	36	10.33	18.5	15.42	284.8	1.32
dconvert	25	6	47	12.47	14.16	4	39	9.83	18.2	13.94	420	1.29
dsystemC	27	4	32	8.36	12.84	3	37	9.83	18.11	12.94	476.9	1.41
glue2	2	14	29	10.61	20.15	9	11	1.41	10	0.92	18.8	0.50
rsd1	6	11	75	24.67	25.57	10	23	4.8	16.33	9.81	94.6	0.64

TABLE IV
RESULTING SUPER-CLUSTERS BASED ON *module*, *always*, AND *if* STATEMENT DECLARATIONS.

Bench-mark	ALWAYS					IF					MODULE				
	blocks	clusters	%Map	STD	Geo	blocks	clusters	%Map	STD	Geo	blocks	clusters	%Map	STD	Geo
cfc18	114:74	166	0.26	18.57	458	114:0	0	0.00	0.00	0	1:1	627	1.00	0.00	49
cft8	387:91	285	0.28	56.27	763	387:2	10	0.01	1.17	7	1:1	911	0.88	0.32	63
desa	11:1	32	0.12	2.87	25	0:0	0	0.00	0.00	0	11:3	202	0.76	4.12	71
iir1	3:2	21	0.15	2.12	23	3:2	21	0.15	2.12	23	5:5	142	1.00	2.62	74
oc54	22:18	399	0.75	9.34	364	26:15	171	0.32	11.80	302	8:8	464	0.88	3.80	184
pajf	7:5	166	1.00	1.34	65	13:3	85	0.51	1.34	41	4:3	166	1.00	0.52	34
rsd2	42:22	264	0.53	5.19	237	17:17	143	0.29	6.74	173	23:7	267	0.54	4.68	175
dconvert	1:1	300	0.67	0.32	41	1:1	300	0.67	0.32	41	1:1	445	1.00	0.00	41
dsystemC	1:1	289	0.71	0.00	39	1:1	289	0.71	0.00	39	1:1	409	1.00	0.00	39
glue2	4:3	41	0.95	1.05	16	13:6	40	0.93	0.48	12	1:1	41	0.95	0.48	12
rsd1	54:25	174	0.91	8.44	155	21:20	121	0.63	6.43	102	31:7	180	0.94	3.84	110

relate to the final placements generated by the SA placer (for ten random seeds). The results for all three constructs are reported using the same format. Subcolumn 1 states the total number of each construct located in the source HDL versus the total number that we could successfully map to clusters on the final design. The number of constructs that could be successfully mapped to clusters is less than the total as certain (unknown) logic optimizations in RASP remove high-level naming information from the output netlist, meaning that these high-level constructs are "lost" in the final placement. Therefore, the number and percentage of clusters successfully mapped to super-clusters are listed in Subcolumns 2 and 3 respectively. Finally, Subcolumns 4 and 5 list the standard deviation and geometric mean of the number clusters per HDL construct-based super-cluster.

C. Data Analysis

An examination of the results shown in Table III indicates that AFFIN2 generally creates fewer super-clusters (excluding *pajf*) with a more even distribution of clusters and a larger number of clusters per super-cluster on average. In contrast, the Local MD of super-clusters in AFFIN2 is also generally larger than that of AFFIN1, which may suggest poorer locality of AFFIN2 super-clusters in the SA placement. However, Figure 4 highlights how the average Local MD per cluster for AFFIN 2 is generally less than AFFIN1. A smaller value for the MD of a super-cluster means that the clusters within it are more tightly packed in their placement on the FPGA.

The correlation between HDL constructs and final placements is harder to interpret as shown in Table IV. First of all, the loss of clusters due to loss of naming information during logic optimization (as discussed in Section VI-B) results in only "samples" of these HDL-construct-based super-clusters being visible for analysis. Furthermore, in the cases of the

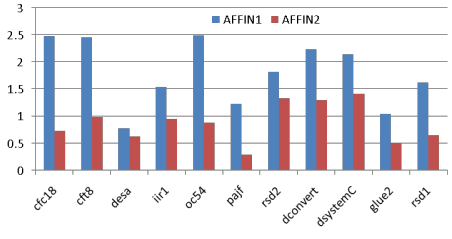


Fig. 4. Average local MD per cluster for AFFIN1 and AFFIN2

always and *if* blocks, these are generally small percentages of the design; in fact, only the bottom four entries of Table IV can provide super-cluster mappings that represent at least 60% of the design. Of these four benchmarks, *dconvert* and *dsystemC* highlight an important consideration for using HDL-constructs to generate super-clusters: they only have one *module*, *always*, and *if* block that encapsulates most of the clusters in the design. Therefore, using HDL-constructs to generate a coarse grain mapping of the design requires that the design be large enough and that the designer has used these constructs to encapsulate sufficient design hierarchy (i.e., *modules*) and/or control structures (i.e., *always*, *if*) that multiple meaningful super-clusters can be generated.

VII. Conclusion and Future Work

In this paper, we have summarized a study of how high-level information may be reconstructed by SA in its final placement. We presented a method of evaluating the locality of coarse grain substructures within a placement along with an analysis of how specific types of high-level information correlate to an SA's final placement. Our results suggest that high-level information can be used to find coarse grain mappings that could improve placement algorithms. Finally, we described and implemented two different versions of the affinity propagation algorithm and demonstrated how the similarity factors (AFFIN2) can be tuned to obtain better high-level information with a better relation to their final placement on an FPGA.

Currently, we are looking into methods of ensuring that *all* design-level information is maintained throughout the entire CAD flow. We are also examining other possible structures (i.e., Buses and Datapath) for possible correlations in the final placement to determine which structure(s) has the best correlation in the final placement. Future work includes investigating new possible similarity factors that may result in coarse grain structures with better locality on a SA placement.

References

[1] W. E. Donath, "Complexity theory and design automation," in *In Proc. of the 17th Design Automation Conference*, 1980, pp. 412–419.

[2] A. Ludwin, V. Betz, and K. Padalia, "High-quality, deterministic parallel placement for fpgas on commodity hardware," in *Proc. of the 16th Int'l ACM/SIGDA Symp. on FPGAs*, 2008, pp. 14–23.

[3] B. J. Frey and D. Dueck, "Clustering by passing messages between data points,," in *Science*, vol. 315, Feb. 2007, pp. 972–976.

[4] B. M. Riess and G. G. Ettelt, "Speed: Fast and Efficient Timing Driven Placement," in *IEEE Int'l Symp. on Circuits*, 1995, pp. 377–380.

[5] K. Vorwerk, A. Kennings, and A. Vannelli, "Engineering details of a stable force-directed placer," in *Proc. of the 2004 IEEE/ACM Int'l Conf on CAD*, 2004, pp. 573–580.

[6] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983.

[7] V. Betz, J. Rose, and A. Marquardt, *Architecture and CAD for Deep-Submicron FPGAs*. Kluwer Academic Publishers, 1999.

[8] Y. Sankar and J. Rose, "Trading quality for compile time: ultra-fast placement for fpgas," in *ACM/SIGDA Int'l Symp. on FPGAs*, 1999, pp. 157–166.

[9] J. M. Varanelli and J. P. Cohoon, "A fast method for generalized starting temperature determination in homogeneous two-stage simulated annealing systems," *Computers and Operations Research*, vol. 26, no. 5, pp. 481–503, 1999.

[10] M. Haldar, A. Nayak, A. Choudhary, and P. Banerjee, "Parallel algorithms for fpga placement," in *10th Great Lakes Symp. on VLSI*, 2000, pp. 86–94.

[11] D. Behrens, K. Harbich, and E. Barke, "Circuit partitioning using high-level design information," in *Conference on Integrated Design - process Technology*, 1996, pp. 259–266.

[12] Y. W. Tsay, W. J. Fang, A. C. Wu, and Y. L. Lin, "Preserving hdl synthesis hierarchy for cell placement," *Proc. of the 1997 Int'l symp. on Physical design*, pp. 169–174, 1997.

[13] J. M. Emmert and D. Bhatia, "A methodology for fast fpga floorplanning," in *Proc. of the 7th Int'l ACM/SIGDA Symp. on FPGAs*, February 21-23, 1999, pp. 47–56.

[14] R. Tessier, "Fast placement approaches for fpgas," *ACM Trans. on Design Automation of Electronic Systems*, vol. 7, no. 2, pp. 284–305, April 2002.

[15] L. K. Grover, "Standard cell placement using simulated sintering," in *Proc. of the 24th ACM/IEEE DAC*, 1987, pp. 56–59.

[16] P. Jamieson and J. Rose, "A Verilog RTL Synthesis Tool for Heterogeneous FPGAs," in *Field-Programmable Logic and Applications*, 2005, pp. 305–310.

[17] S. Williams, "ICARUS Verilog at http://www.icarus.com/eda/verilog/," 2007.

[18] P. A. Jamieson, K. B. Kent, F. Gharibian, and L. Shannon, "Odin II - An Open-source Verilog HDL Synthesis Tool for Academic CAD Flows," in *Field-Programmable Custom Computing Machines, Annual IEEE Symp. on*, 2010.

[19] J. Jones, "Abstract syntax tree implementation idioms," in *Proceedings of the 10th Conference on Pattern Languages of Programs (PLoP2003)*, 2003.

[20] J. Cong, J. Peck, and Y. Ding, "RASP: A General Logic Synthesis System for SRAM-Based FPGAs," in *ACM/SIGDA Int'l Symp. on FPGAs*, 1996, pp. 137–143.

[21] A. Marquardt, V. Betz, and J. Rose, "Using Cluster-Based Logic Blocks and Timing-Driven Packing to Improve FPGA Speed and Density," in *ACM/SIGDA Int'l Symp. on FPGAs*, Monterey, CA, 1999, pp. 37–46.

[22] J. Luu, I. Kuon, P. Jamieson, T. Campbell, A. Ye, W. M. Fang, and J. Rose, "VPR 5.0: FPGA CAD and Architecture xploration Tools with Single-Driver Routing, Heterogeneity and Process Scaling," in *ACM/SIGDA Int'l Symp. on FPGAs*, Feb 2009.

Farnaz Gharibian received her M.Sc. degree from University of New Brunswick, Fredericton, Canada in 2008. She is a Ph.D. student in the School of Engineering Science at Simon Fraser University, working with Dr. Lesley Shannon.

A Micro Electromagnetic Generator for Vibration Energy Harvesting

Alireza Hekmati, *Member, IEEE*

***Abstract*—This paper presents two prototypes of a miniaturized electromagnetic generator which consist of two moving magnets within a cylindrical coil. In prototype A, the magnetization of the moving magnets is in the same direction, so the magnets can be stuck together easily. In prototype B, the magnetization of the moving magnets is in the opposite direction, so a cylindrical soft magnetic spacer with the same radius as the cylindrical moving magnets is needed to separate moving magnets to reduce the repulsion force between them. The simulation has been done with Comsol Multiphysics software, which is a powerful tool to demonstrate the pattern of magnetic field and calculate the induced current in the coil. The simulation results demonstrate that in prototype B, the magnetic field is better concentrated on the coil area. Therefore, this prototype produces more induced current in the coil, and hence, we can expect more power from it. Finally, prototype A has been built and tested by a shaker within the frequency range of 5 "Hz" to 50 "Hz".**

***Index terms*—average power, energy harvesting, magnetization, micro electromagnetic generator.**

I. INTRODUCTION

Energy harvesting refers to all the methods which enable us to convert diverse forms of ambient energy into the storable form of electrical energy. The rapidly decreasing size, cost, and power consumption of sensors and electronics has opened up the relatively new research field of energy harvesting. The goal is to harvest enough ambient energy to power a standalone sensor and/or actuator system.

The author is with the Department of Mechatronic Engineering, Simon Fraser University, Surrey, BC V3T 0A3, Canada (e-mail: alireza_hekmati@sfu.ca).

On the other hand, wireless sensor systems are receiving increasing interest since the use of them offers several advantages over existing, wired methodologies. Factors include flexibility, ease of implementation, and the ability to facilitate the placement of sensors in previously inaccessible locations. The ability to retrofit systems without having to consider issues such as cabling, offers a significant advantage in some applications. For example, in condition-based monitoring (CBM), embedded wireless micro sensors can provide continuous monitoring of machine and structural health without the expense and inconvenience of including wiring looms [1]. The wires (and associated connectors) are often a source of failure in such systems and present a considerable cost issue.

At present, many wireless sensor nodes are battery powered and operate on an extremely economical energy budget, since continuous battery replacement is not compatible with embedded applications nor is it feasible for networks with large numbers of nodes [2]. The advances made in low power wireless systems present an opportunity for alternative types of power source rather than traditional batteries. Solutions such as micro fuel cells [3] and micro turbine generators [4] are capable of high levels of energy and power density. However, they involve the use of chemical energy and require re-fuelling when their supplies are exhausted.

The sources of energy available for harvesting are essentially of four forms: light, radio-frequency (RF) electromagnetic radiation, thermal gradients, and motion, including fluid flow [5]–[8].

Kinetic energy generators or motion generators convert the mechanical movement energy, which present in the application environment, into the electrical energy. Kinetic energy is typically present in the form of vibrations, random displacements, or forces, and it is typically converted into electrical energy using electromagnetic, piezoelectric, or electrostatic mechanisms. Suitable vibrations can be found in numerous applications including common household goods (fridges, washing machines, microwave ovens etc.), industrial plant equipment, moving structures such as

automobiles and airplanes, and structures such as buildings and bridges [9]. The amount of energy generated by this approach depends fundamentally upon the quantity and form of the kinetic energy available in the application environment and the efficiency of the generator and the power conversion electronics.

Electromagnetic induction, first discovered by Faraday in 1831, involves the generation of electric current in a conductor located within a magnetic field. The conductor typically takes the form of a coil and the electricity is generated by either the relative movement of the magnet and coil, or because of changes in the magnetic field. In the former case, the amount of electricity generated depends upon the strength of the magnetic field, the velocity of the relative motion and the number of turns of the coil [10]. Vibration-based micro-power generators have a key design. They are optimally tuned to harvest energy within a narrow frequency band around the oscillator's natural frequency. Outside this band, the output power is too low to be conditioned and utilized.

II. BASIC THEORY

Resonant generators can be modeled as a second-order, spring–mass–damper system with base excitation [9]. Figure 1 presents a general example of such a system based on a seismic mass, *m*, a spring of stiffness, *k*, and total energy losses within the system, c_T, which consist of parasitic loss mechanisms (e.g. air damping), represented by c_P, and electrical energy extracted by the transduction mechanism, represented by c_E.

These generators are intended to operate at their resonant frequency, and for optimum energy extraction, they should be designed such that their resonant frequency coincides with the vibrations present in the intended application environment [7]. Assuming that the generator is driven by a harmonic base excitation y (t) = Y sin (ωt), which will result in a net displacement, z(t), between the mass and the frame. The average power dissipated within the damper (i.e. the power extracted by the transduction mechanism and the power lost through parasitic damping mechanisms) is given by [10]:

$$P_{av} = \frac{m\left(\frac{\omega}{\omega_n}\right)^3 \xi_T Y^2 \omega^3}{\left\{\left[1-\left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left(2\xi_T \frac{\omega}{\omega_n}\right)^2\right\}} \quad (1)$$

where ζ_T is the total damping ratio given by $\zeta_T = c_T /2m\omega_n$. When the device is operated at ω_n, the maximum power dissipation within the generator occurs, and in this case, P_{av} is given by [7]:

$$P = \frac{mY^2\omega_n^3}{4\zeta_T} \quad (2)$$

Incorporating the parasitic damping and electrical damping into (2) gives the power delivered to the electrical load [9]:

$$P_L = \frac{m\xi_E Y^2 \omega_n^3}{4(\xi_P + \xi_E)^2} \quad (3)$$

From the above equation, obviously the maximum power is delivered to the load when $\zeta_E = \zeta_P$ (i.e. damping arising from the electrical domain should equal mechanical losses). In this case, (3) simplifies to:

$$P_L = \frac{mY^2\omega^3}{16\xi_P} \quad (4)$$

The coil resistance and the load resistance affect the damping factor arising from electromagnetic transduction, c_E, which can be estimated from [9]:

$$C_E = \frac{(NlB)^2}{R_L + R_{coil} + j\omega L_{coil}} \quad (5)$$

where *N* is the number of turns in the generator coil, *l* is the side length of the coil, and *B* is the flux density to which the coil is subjected. R_L, R_{Coil}, and L_{Coil} are the load resistance, coil resistance, and coil inductance respectively.

Equation (5) demonstrates that R_L can be used to adjust c_E to match c_P to maximize the power delivered to the load. It can be proved that the optimum R_L and the maximum average power delivered to the load can be found from (6) and (7), respectively [9].

$$R_{L_{Opt}} = R_{coil} + \frac{(NlB)^2}{C_P}(1) \quad (6)$$

$$P_{L_{Max}} = \frac{m\omega_n^3 Y^2}{16\zeta_P}\left(1 - \frac{R_{coil}}{R_{load}}\right) \quad (7)$$

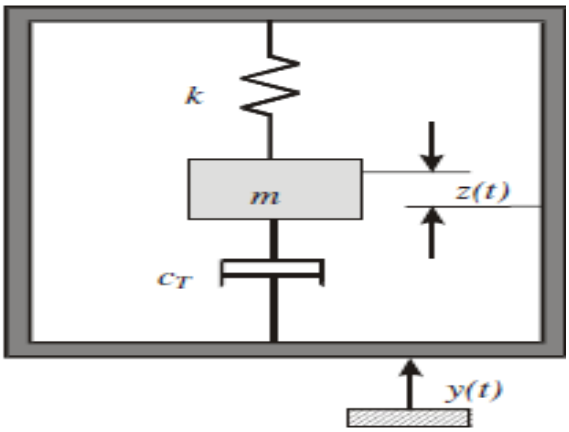


Fig. 1. Model of a linear, inertial generator.

III. STRUCTURE OF THE GENERATOR

The configuration of the magnetic generator structure is presented in figure 2. The basic idea of this structure is that axially magnetized permanent magnets are placed vertically inside a tube, and a coil is wrapped around the moving magnets. During vibration, the relative motion between magnets and the coil leads to a varying magnetic flux through the coil windings. Consequently, a voltage proportional to the rate of flux change is induced in the coil [10].

The structure comprises of two cylindrical moving magnets, which the radius and the height of each one is equal to 1.585 mm and 3.17 mm, respectively. The coil's shape is also cylindrical with the inner radius of 2.015 mm, the outer radius of 4.220 mm, and the height of 13.550 mm.

A magnet is fixed at one end of the generator to suspend moving magnets inside the coil due to the repulsive force. Therefore, the moving magnets will not contact to the ends of tube, and as a result of that, the power dissipation due to the friction will decrease. There is the possibility to use a spring to suspend moving magnets, but this method creates more space for magnets to move and hence more voltage will be induced in the coil.

In prototype A, the magnetization direction of the moving magnets are the same, but in prototype B, because the magnetization direction of the moving magnets are opposing, moving magnets have been separated by soft magnetic spacer to reduce the repulsion force between two magnets.

Assuming motion in the *z*-direction, the voltage induced in the coil is given by:

$$V_{Coil} = N\frac{d\Phi}{dt} = N\left(\frac{d\Phi}{dz}\right)\left(\frac{dz}{dt}\right) \qquad (8)$$

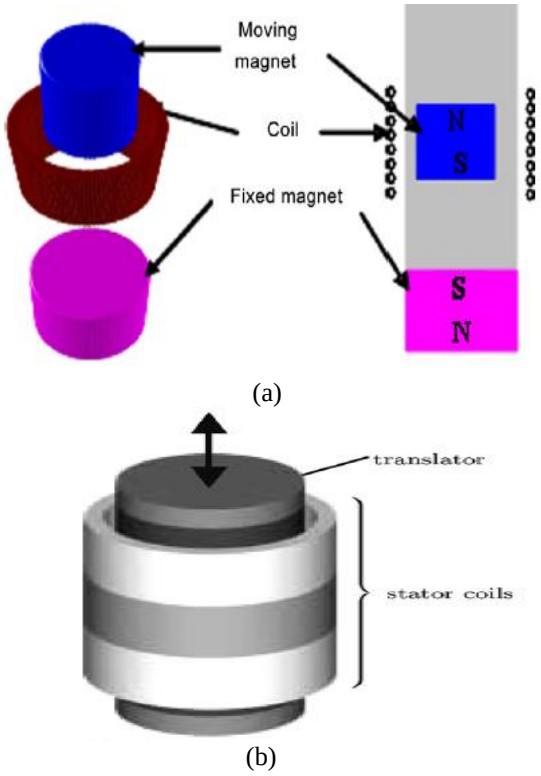


Fig. 2. (a) magnetic generator structure. (b) 3D model.

where *N* is the number of turns in the coil, Φ is the average turn flux linkage, and (d*z*/d*t*) is the velocity of movement in the *z*-direction. From this equation, the important parameters for the generation of voltage are the flux linkage gradient, and displacement or velocity [11].

The used generator architecture has the following advantageous properties:

• *Moving magnet*: Compared to moving coil generators, the selected moving magnet architecture avoids electrical connections to moving parts.

• *Rotation-symmetric design*: The generator topology is rotation-symmetric, which generally simplifies simulation (i.e. 2D simulation possible) and construction. Moreover, this structure can be easily built since the generator simply consists of magnets and coil without the need for any mechanical beam.

IV. SIMULATION ANALYSIS

A. *Prototype A*

As mentioned before, in this prototype, the magnetization direction of the moving magnets are the same. The generator structure has been modeled, using finite element analysis (FEA) with the Comsol Multiphysics software, in order to understand the flux

linkage with the coil, and to calculate the induced current and the average power in the coil.

In simulation process, the magnetization has been chosen to be just in z-direction with the value of 1.19 “MA/m”, and the material of the coil is copper. The excitation of the structure is defined to be as Z = 0.002 sin (20.π.t) “m”.

The results of an axial-symmetric finite element simulation of the prototype A, which demonstrates the magnetic field lines and the surface induced current density, is presented in figure 3. The integral result of induced current density over the whole surface of the coil is presented in figure 4, so it demonstrates the produced current in the coil due to relative motion between magnets and the coil. Also, this induced current in the coil in the frequency domain is demonstrated in figure 5. Finally, the integral result of power flow over the whole surface of the coil is presented in figure 6, so it determines the produced average power. Analyzing this figure demonstrates that the maximum power, which can be extracted from prototype A, is 1.553 “mW”.

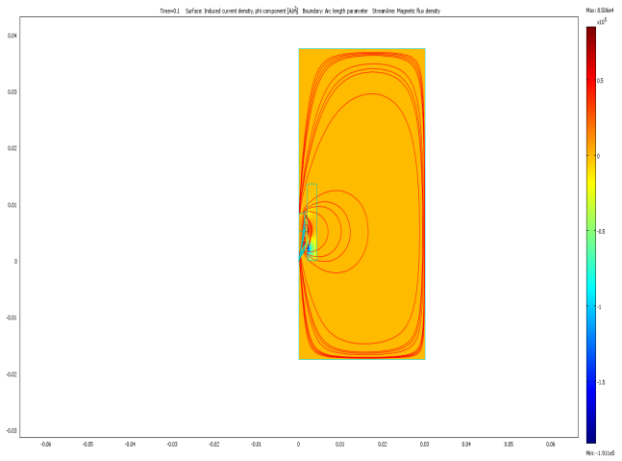


Fig. 3. Magnetic flux density and induced current density of prototype A.

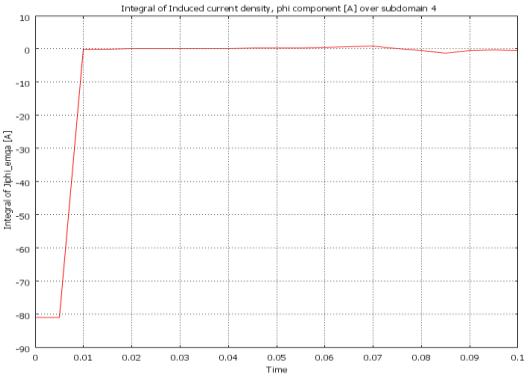


Fig. 4. Total induced current in the coil.

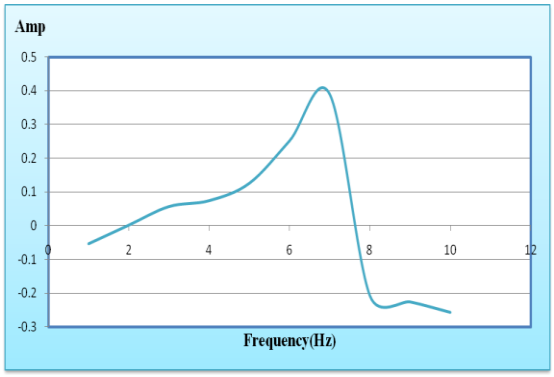


Fig. 5. Frequency response of induced current for prototype A.

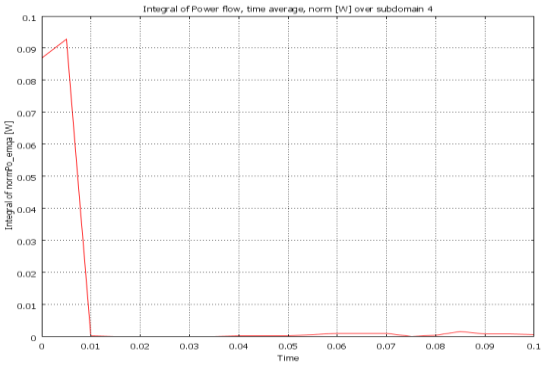


Fig. 6. The average power which can be extracted from prototype A.

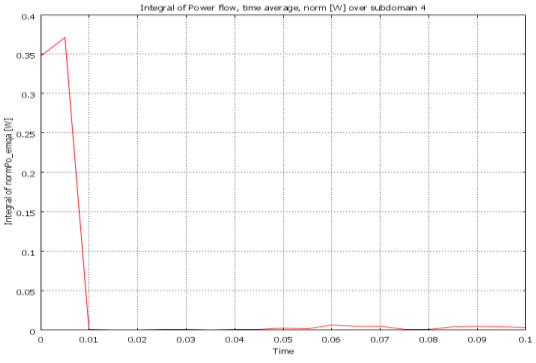


Fig. 7. The effect of changing the magnetization of moving magnets on average power in prototype A.

Here is represented the effect of changing five different parameters on maximum average power with the aim of simulation. These parameters are: magnetization of magnet, material of the coil, length of the coil, amplitude level, and frequency.

1) Magnetization of magnet: Here the magnetization of moving magnets is changed from 1.19 “MA/m” to 2.38

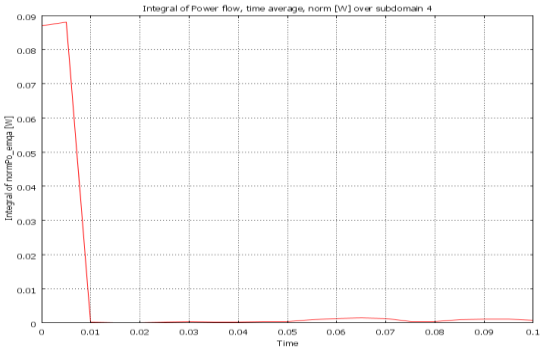


Fig. 8. The effect of changing the material of the coil on average power in prototype A.

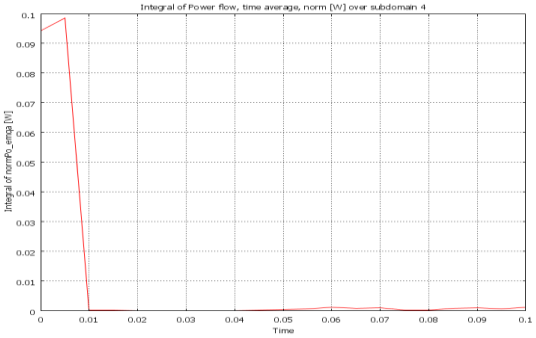


Fig. 9. The effect of changing the length of the coil on average power in prototype A.

"MA/m" (i.e. stronger magnets). The result of this change is presented in figure 7, and from this figure, it can be observed that the maximum average power is equal to 6.672 "mW", which is much more than the primary case.

2) Material of the coil: Here the material of the coil is changed from Copper to Aluminum (i.e. increasing the value of R_{Coil}). The affect of this change is demonstrated in figure 8, and from this figure, it can be comprehended that the maximum average power is equal to 1.517 "mW", which is a little lower than the primary case.

3) Length of the coil: Here the length of the coil is changed from 13.55 "mm" to 15.55 "mm" (i.e. higher length for the coil), and the result of this change is presented in figure 9. From this figure, the maximum average power is equal to 1.224 "mW", which is lower than the primary case, since with increasing the length of the coil, R_{Coil} is increasing, too.

4) Amplitude level: Here the amplitude level is changed from 2 "mm" to 1 "mm" (i.e. lower amplitude), and the influence of changing this parameter on the maximum average power is demonstrated in Figure 10. From this figure it can be understood that the maximum average power is equal to 0.9328 "mW", which is much lower than the primary case.

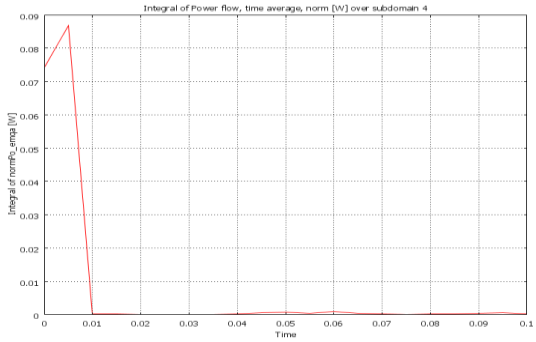


Fig. 10. The effect of changing the amplitude level on average power in prototype A.

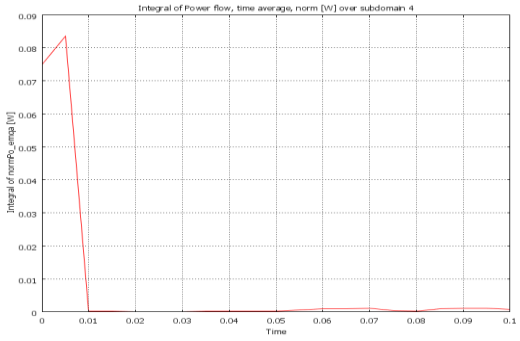


Fig. 11. The effect of changing the frequency on average power in prototype A.

5) Frequency: Here the frequency of moving magnets is changed from 10 "Hz" to 20 "Hz" (i.e. higher frequency), and the impact of this change is presented in figure 11. From this figure, it can be concluded that the maximum average power is equal to 1.161 "mW", which is lower than the primary case, since in this case, the system is not vibrating on its natural frequency.

At the end of this part, it can be concluded that the following factors have the most important role to design a high efficiency micro electromagnetic generator:

- Stronger magnets;
- Lower impedance of the coil;
- Shorter length of the coil;
- Higher input amplitude.

Finally, the generator has to be designed in such a way that it resonates at its natural frequency.

B. *Prototype B*

As mentioned before, in this prototype, the magnetization direction of moving magnets are opposing. Moreover, in this prototype, a cylindrical soft magnetic spacer with the same radius as cylindrical moving magnets and the height of 1.585 "mm" is separated two moving magnets to reduce the repulsion force between them. All the set up and steps are the same as prototype A, and the results were presented in the following figures:

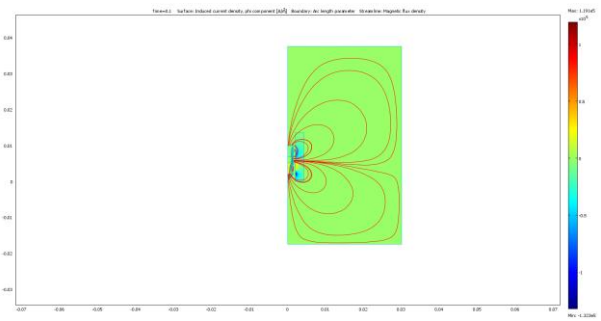


Fig. 12. Magnetic flux density and induced current density of prototype B.

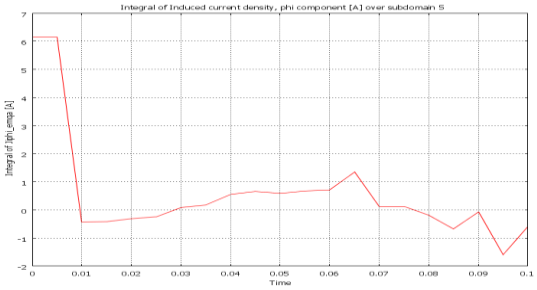


Fig. 13. Total induced current in the coil.

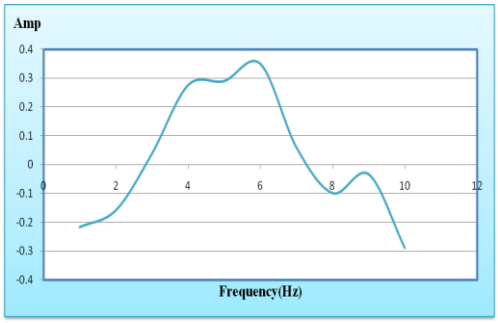


Fig. 14. Frequency response of induced current for prototype B.

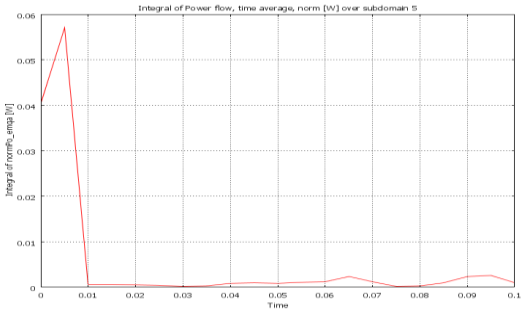


Fig. 15. The average power which can be extracted from prototype B.

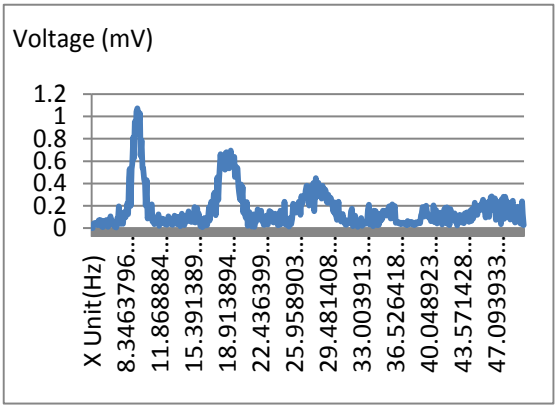


Fig. 16. Voltage versus frequency in experimental result for prototype A.

Analyzing figure 15 demonstrates that the maximum power which can be extracted from prototype B is 2.626 “mW”.

V. EXPERIMENTAL RESULT

Testing the prototype A was conducted, while using a shaker unit with accelerometer feedback. The following result was taken at the amplitude level of 2 “mm” and the frequency range of 5 “Hz” to 50 “Hz”.

The maximum voltage peak in figure 16 has been occurred in the frequency equal to 10 “Hz”. Therefore, 10 “Hz” is the resonance frequency of the system. Consequently, the simulations were conducted at this frequency. From figure 16, the maximum induced voltage is equal to 1.1 “mV”. Thus, according to equation $P = \frac{V^2}{R}$ and considering the fact that R_{Coil} is equal to 0.5 “Ω”, the power is equal to 0.00242 “mW”, which is far away from what has been expected from simulation results. One of the reasons for this difference between simulation and experimental result is that in experimental setup, the moving magnets directly have a contact with one end of the tube surface, which obviously increases the damping loss.

VI. CONCLUSION

The idea behind this study was to develop a suitable electromagnetic energy harvester for body motion, which could supply power to body worn sensors or electronic devices. This paper introduced the structure of two prototypes of micro electromagnetic generator.

Simulation results demonstrated that prototype A, which has the same magnetization direction of moving magnets, is able to produce the maximum average power equal to 1.553 “mW”, but prototype B, which has the opposite magnetization direction of moving magnets, is able to produce the maximum average power of 2.626 “mW”. The main reason of higher efficiency for prototype B is because of better concentration of magnetic flux in the coil area.

Prototype A has been fabricated and tested on the shaker, but the experimental result was far away from simulation result. One of the reasons for this difference between simulation and experimental result is that in experimental setup, the moving magnets directly have a contact with one end of the tube surface, which obviously increases the damping loss.

The average generated electrical power of the vibrational generator is proportional to cube of the frequency. Therefore, one way to enhance the average power is either increasing the total stiffness of the structure or decreasing the total mass of it, to increase the natural frequency of the system. Moreover, the average power is proportional to square of acceleration level. So, with higher acceleration level, the generator will be able to produce more average power. Finally, further research is needed regarding energy storage and reliability.

REFERENCES

[1] Bult K *et al* 1996 A distributed, wireless MEMS technology for condition based maintenance *Proc. 1996 Integrated Monitoring, Diagnostics and Failure Prevention Conference, Society of Machine Failure Protection Technology (MPFT) (Mobile, AL, USA).*

[2] Raghunathan V, Schurgers C, Park S and Skrivastava M B 2002 Energy-aware wireless microsensor networks *IEEE Signal Process. Mag.* .

[3] Banazwski B and Shah R K 2003 The role of fuel cells for consumer electronic products and toys *Proc. 1st Int. Conf. on Fuel Cell Science, Engineering and Technology (Rochester, NY).*

[4] Epstein A H 2004 Millimeter-scale, micro-electro-mechanical systems gas turbine engines *J. Eng. Gas Turbines Power.*

[5] Koeneman P B, Busche-Vishniac I J and Wood K L 1997 Feasibility of micro power supplies for MEMS *IEEE J. Microelectomech. Syst.*

[6] G¨orge G, Kirstein M and Erbel R 2001 Microgenerators for energy autarkic pacemakers and defibrillators: fact or fiction *Herz.*

[7] Amirtharajah R and Chandrakasan A P 1998 Self-powered signal processing using vibration-based power generation *IEEE J. Solid State Circuits.*

[8] Jacobson S A and Epstein A H 2003 An informal survey of power MEMS *Proc. Int. Symp. on Micro-Mechanical Engineering ISMME (Japan).*

[9] Roundy S, Wright P K and Rabaye J 2003 A study of low level vibrations as a power source for wireless sensor nodes *Comput. Commun.*

[10] Beeby S P, Tudor M J and White N M 2006 Energy harvesting vibration sources for Microsystems applications, School of Electronics and Computer Science, University of Southampton.

[11] Starner T and Paradiso J A 2004 Human generated power for mobile electronics *Low Power Electronics Design,* ed C Piguet (Boca Raton, FL: CRC Press)

Alireza Hekmati received the B.Sc. degree in electrical engineering in 2009 from Iran University of Science and Technology, Tehran, Iran. He is currently studying Master of Applied Sciences in mechatronic systems engineering at Simon Fraser University, Vancouver, BC.

He joined the Petropars Company, Tehran, Iran, in 2008, where he directed work on diverse types of sensors to detect smoke, gas, and fire in industrial environments. He is currently researching on diverse kinds of energy harvesters.

A Structural Equation Model for Cross-Industry Customer Satisfaction in Iran

Zhila Pirmoradi [1], Monire Jalili [2]

[1] Department of Mechatronic Systems Engineering, Simon Fraser University, Vancouver, V3T 0A3, Canada

[2] Department of Industrial Engineering & Management Systems, University of Central Florida, Florida 32816, USA

Email: zpirmora@sfu.ca, mjalili@knights.ucf.edu

***Abstract*- Identifying and recognizing the factors that affect customer satisfaction is of remarkable importance for increasing his/her satisfaction. A number of such factors do not depend on product/service type. In other words, regardless of what the product/service is, the manufacturer/service provider should care about a number of factors that affect customers' perception and judgment about the product. In this paper, we will develop a causal model containing general factors affecting satisfaction and loyalty of customers in Iran. To build our model, we have used the most common effective factors which result in satisfaction according to previous research done about various products in different countries. Based on data gathered through questionnaires, customers' perceptions and judgments about such factors are gathered for a number of industries, and the model fitness is then tested through LISREL 8.3. The analysis results confirm fitness of the proposed model to the data set and show that most of the measurement variables predict the underlying factor properly.**

Keywords: *causal model, customer satisfaction, structural equation models*

I. INTRODUCTION

Satisfaction is defined as after-consumption assessments and perceptions of customer about a product or service [1]. As a conclusion of previous research, satisfaction of the customers serves as the key to long-term success of the businesses [2] and many of the managers and business owners consider satisfaction as a proper indicator for their financial performance nowadays [3]. Satisfaction and loyalty are of exploratory nature and therefore several approaches may be applicable to assess them. *Expectancy-disconfirmation* approach, *performance-only* approach, and the *service quality versus the service satisfaction* approach are some examples [4]. During the recent decades, efforts to identify customer's needs have been one of the main concerns of industries and service providers [5]. Increasing rate of loyalty, more tolerance against price changes, and increased reputation for the organization are some of the benefits of satisfied customers [6]. Another reason for striving for more satisfied customers is that the costs of acquiring new customers are considerably more than required budget for keeping current ones [7].

Researchers have been interested in finding a way to compare the customers' level of satisfaction across different industries, or industries of the same kind at national levels. For measuring level of satisfaction, a number of countries have developed indices which are applied to consuming goods and services of that country [8]. These indices (we will call them as CSIs from now), aim to measure level of satisfaction and can serve as a tool for monitoring the progress of companies performance, and for identifying the differences between performance of businesses with the same type of activity. Most of these indices are models for predicting customer satisfaction by means of introducing a number of factors that are believed to affect satisfaction in cause-and-effect relations (e.g., customer expectations, perceived value, etc.).

The American Customer Satisfaction Index (ACSI) as one of the most popular indices was developed mainly for measuring and predicting customer loyalty, based on the direct relation discovered between satisfaction and loyalty [9]. Customer expectation, perceived service quality and perceived value are antecedents of satisfaction in ACSI, and customer complaints, and loyalty are satisfaction consequences in this model [10]. According to research, the CSIs can be used for predicting the firms' profitability and their share of market [11]. European CSI is another similar index widely used in Europe [4].

Structural Equation Models or SEM are causal models applied to evaluate the factors affecting satisfaction. This kind of modeling allows measuring unobservable variables (or latent variables, or constructs, or factors) by the means of observable or manifest variables or indicators [12]. SEM is a proper technique for the customer research field due to its abilities for taking into account the measurement error, the structural error, and the correlations among every pair of factors. This technique provides the possibility of estimating all assumed relationships among the variables, and allows testing the importance of each factor [13]. In predictive models, for any given value of indicator (or predictor) variables, the conditional expected values of dependent variables can be estimated based on a series of equations in

which the dependent or unobserved variables are supposed to be linear combinations of the predictive variables [7]. Estimation approaches such as Maximum Likelihood (MLE) or Partial Least Squares (PLS) can be applied to estimate free parameters of the models and then to test different hypotheses about causal relations among the model constructs. MLE, which was developed by Jöreskog in 1973 is a popular approach used in SEM. However, this approach has a number of weaknesses such as necessity of large sample sizes, necessity of interval scaling for the data, and necessity of assumptions about multivariate normality of the variables. Since such constraints do not appear in PLS, the recent approach is preferred in cases like ACSI [14].

Developing a valid model for measuring satisfaction in all industries is not an easy task and sometimes it may not be possible to aggregate all the vital factors in a single and simple model [15]. On the other hand, the need for such a common measure is increasing, because it facilitates benchmarking and continuous performance improvement. Such a measure will provide the possibility of comparing one's performance with others' and if it is a worldwide scale, it will provide the decision makers and the business leaders with valuable pieces of information about their business, their competitors and other reputable and well-established firms of the same activity type [4]. As discussed before, the CSIs are well-known examples of such measures. Nicholls et al. (1998) tried to make a comparison between hospitality-oriented businesses and sports-oriented ones; and McDougal and Levesque (2000) studied the impact of parameters such as perceived value on customer satisfaction in four service sectors [4].

Similar to the ACSI and other indices, many of the proposed models focus on developing a CSI to serve as a national cross-industry scale for measuring satisfaction [10]. But, customizing these developments necessitates some modification dependent on the culture and particular characteristics of the nation under study. Therefore, providing any nation with a proper and reliable model requires precise studies about that nation markets, the constraints and policies that run the market, and expectations, preferences, and taste of its customers.

Beside the nations, customization is needed for industries or services as well, but in this paper we aim to focus on identifying basic factors that affect satisfaction regardless of the particular product/service that is offered to the customer. In fact, what we focused on is to provide a guiding tool applicable by almost all business types. For small and newly developed businesses, it might not be effective to follow a complicated process to discover all of the customer's needs and expectations in detail. So, our paper supposes that less developed businesses and novice managers ought to be aware of the core expectations and needs of their customers. However, without providing these fundaments, a company cannot plan for achieving higher market share or even survive.

Some researchers believe that the cross industry measures might be less effective because businesses have different contexts and are not comparable through a single scale. Therefore they suggest that satisfaction should not be measured by the same scale for all industries [5]. Considering all these ideas together, what persuaded us for doing the present study was to provide a standard and unique measure to reflect the core antecedents and consequences of satisfaction, regardless of the industry context. Although using specific measurement criteria for each industry will make the measurement more reliable and precise, particular benefits of such a general model should not be ignored as well. Main incentive for developing such a cross industry model was to aid young businesses in planning for meeting the core needs and expectations of their customers. Naturally, after growing and making more developments, such companies would also seek for more specific measures and would consider satisfaction measurement programs in more detail.

In our paper, a SEM is proposed about the cause and effect relationships among satisfaction antecedents and its consequences. The proposed model is tested through evaluating customers' perceptions and viewpoints about different products and services. Our assumptions are mainly inspired and adopted from ACSI. On the other hand, particular aspects of Iranian markets are considered for providing a reliable model about Iran. The paper is organized as follows: in section 2, the proposed model is explained in detail. In section 3 the empirical study is explained along with the analysis results, in section 4 we provide conclusion and finally in section 5, suggestions for future research are presented.

II. THE PROPOSED MODEL

With the objective mentioned earlier, and by studying the popular CSIs and other developed causal models for satisfaction, we developed the proposed model through combining the factors which we suppose are appropriate for Iranian customers and more applicable to Iranian markets. The model comprises of the following factors:

Antecedents of customer satisfaction

Perceived Quality: According to previous studies, satisfaction results from comparison of what customer expected from the product/service before use, to what he/she conceives about it after use [8]. A number of researchers contend with inserting *customer expectations* as a factor into satisfaction models [16], [17]. They claim that the expectations of the customers don't affect satisfaction in long run and it strongly depends on the time and market conditions [18]. Therefore we also prefer to consider the expectations in an indirect way: considering its effects on perceived quality.

Perceived Value: The tradeoff between price and quality of a product is another important factor in judgment of the customers [7]. This judgment is called perceived value (PV) and based on past research positively affects satisfaction [6]. Besides, this factor is concluded to be directly affected by the

perceived quality [19], [6]. Perceived value can also be defined as the value that the customer believes that has been offered to him by the supplier [20].

Perceived Image: Perceived image (PI) is the third factor that affects satisfaction positively [17]. This factor is defined as overall judgment of customer about the supplier or the brand. This judgment is resulted from various events such as the past purchases or visits with that supplier, or from any other information source such as competitors of that industry, other customers, etc. Perceived image is supposed to affect the level of satisfaction positively [3]. It also positively affects customer judgment about the value offered to him [17]. Hsu [10] provides the idea that conformation of a positive image about trustworthiness of the supplier is one of the invaluable assets for that supplier, because it influences the customer's willingness to retain his connection to that supplier.

Complaint Management (CM): Complaint can be defined as a conflict between the customer and the supplier which can be turned into opportunities for increasing loyalty if it is properly managed by the supplier [6]. Customer satisfaction is shown to decrease complaints [7], [8]. Johnson et al., (1995) claim that if failure and defects are inevitable parts of the products or services, then the firms will be able to create opportunities through defects to strongly satisfy their customers [8]. It is notable that defect is not causing dissatisfaction per se in the sense that most of the customers have accepted that defect might naturally occur. What leads to dissatisfaction is the improper reaction and lack of supplier responsiveness to resolve and correct the defect. Therefore, complaint management is a factor that is believed to influence customers' loyalty positively and is included in our model. A specific assumption of our study is about the causal relationship between complaint management and perceived image; we propose that proper management of the complaints improves the trustworthiness of the supplier. In addition, since the customer judges more optimistically about responsiveness of the supplier when a defect happens; it improves and reinforces the image about the supplier. So, complaint management influences customer loyalty indirectly by improving the perceived image. This factor is one of the exogenous factors of our model which is assumed to result in satisfaction and loyalty in a different structure than ACSI.

Switching Barriers: Another factor that significantly affects loyalty is the switching barriers (SB). It can be defined as all the activities that the supplier can perform in order to retain its customers and preclude them from leaving [21]. These activities may be included in the marketing strategy of the firm or may be unintentional actions which perform as barriers to customer's switching. In addition to activities done by the supplier, natural obstacles such as distance or some external barriers may exist which are not necessarily formed by the supplier, but can be effective in keeping a customer. Examples of intentional activities can be discounts for loyal customers, providing loyal customers with increased level of services, more rapid delivery, and planning for strong social bonds with the customer in order to decrease his/her tendency toward leaving [6]. The switching barriers can be considered from two dimensions: one is the psychological dimension; for example the customer may concern about the level of service he will receive from new supplier in comparison to the level he receives from the current supplier, and he may prefer retaining the current connection and not think of switching [2]. Also, the fear of cutting old bonds and starting a new connection is a psychological barrier. This type is more unwelcoming specifically in the areas such as dentistry, fitness centers, physiotherapy centers and other similar services in which, there are more contacts between customer and personnel [22]. The second dimension of the switching barriers is economical. Examples of this type are actions, time or money required for getting useful information about acceptable new alternatives before switching, or the difficulties that might exist in making precise comparison among the scope of available services of different suppliers, and specially the difficulty of judgment about services which are not received yet (i.e., service quality assessment will not be possible before receiving it) which makes it difficult to make decision about switching [2].

Therefore, in our proposed model, switching barrier is supposed to positively influence loyalty. This effect indicates another fact that is, retention of the customers may not be the direct result of their satisfaction, but can be promoted by some barriers. The same effect also implies that the firms can take the opportunity of creating barriers to switching, in order to retain their customers. In other words, the companies can exploit barriers such as social bonds and special offers to their current customers. It is notable that these barriers should not be used as a substitute for what actually results in satisfaction. The reason is that while such barriers might create value for the supplier in short term, if a customer is obliged to continue a relation without being truly satisfied but merely because of barriers, he would not avoid any kind of nagging and spreading negative word of mouth [23]. Particularly, we are interested in testing the positive effect of satisfaction on switching barriers. In other words, we will test the causal relation between satisfaction and switching barriers to find if satisfaction can perform as a positive barrier for keeping the customer loyal which implies indirect effect of satisfaction on loyalty through positive barriers.

Consequences of satisfaction

Customer Satisfaction: Satisfaction (CS) is the overall judgment of the customer about the product, the supplier or the sales personnel [24], and the purchasing experience [25], [26]. In whole, it is the result of making a comparison between expectations and perceived performance of a product [5]. This factor is used as the indicator of the satisfaction level in most of the CSIs and therefore, is included in our study as well.

Customer Loyalty: The last included factor in our model is loyalty (LO) which is defined as the long term commitment of a customer to the supplier, and his willingness to retain the connection to that supplier [27]. A loyal customer uses

positive word of mouth and recommends the product and the supplier to others. Price tolerance is another proof of loyalty which means that, a truly satisfied customer is likely to willingly pay more for his desired product in comparison to unsatisfied customers [3]. Repurchasing behavior can be an indicator of loyalty as well. However, it is shown by some researchers that tendency to repurchase may not always result in repurchasing [3].

The model along with the indicators for each factor is illustrated in figure 1. The hypotheses that will be tested are as follows:

H_1: Complaint management positively influences perceived image of the customer.
H_2: Perceived image affects perceived value positively; in other words by improving the image of the organization for a customer, he believes the organization cares more about creating value for him.
H_3: Complaint management directly and positively improves judgment of customer about the value created for him by the supplier.
H_4: Perceived quality has a positive relationship with perceived value which means better judgment about quality and performance of a product increases the created value perceived by the customer.
H_5: The customer's judgment about the value created for him positively influences his level of satisfaction
H_6: Perceived quality positively affects the satisfaction level.
H_7: Customer satisfaction plays a positive role in preventing the customer from switching, and satisfaction is resulting in formation of positive psychological barriers for switching.
H_8: By changing the level of satisfaction, loyalty rate of a customer positively changes and increased satisfaction results in more loyalty.
H_{9a}: If there is a positive relationship between switching barriers and the customer loyalty, it can be concluded that the barriers have been successful in retaining the customer.
H_{9b}: If there is a negative relationship between switching barriers and the customer loyalty, it is conceivable that in spite of some barriers, the customer is leaving and high levels of dissatisfaction can be inferred.
The proposed model is shown in form of a SEM in figure 2 which includes all of the factors which will be estimated in Analysis stage.

III. EMPIRICAL STUDY AND RESULTS

A. Sampling and data

At this stage, proper sampling is required and a proper sample size should be selected to satisfy the desired significance level. We selected 95% confidence interval for our study. Since SEM needs large sample sizes (about 200) [28], and also based on selected significance level for our study the initial sample size chosen was 400, but only 273 of the filled questionnaires were eligible to be considered and relied upon for doing the analysis.

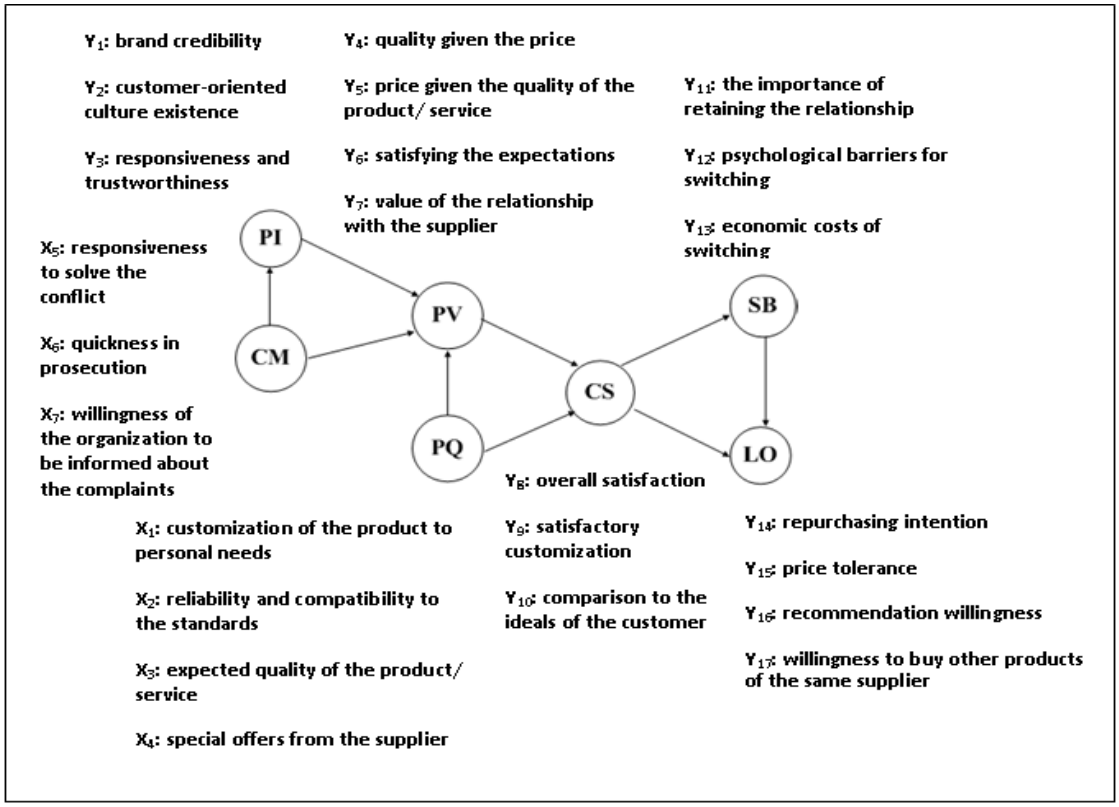


Figure1. Proposed model for cross industry satisfaction

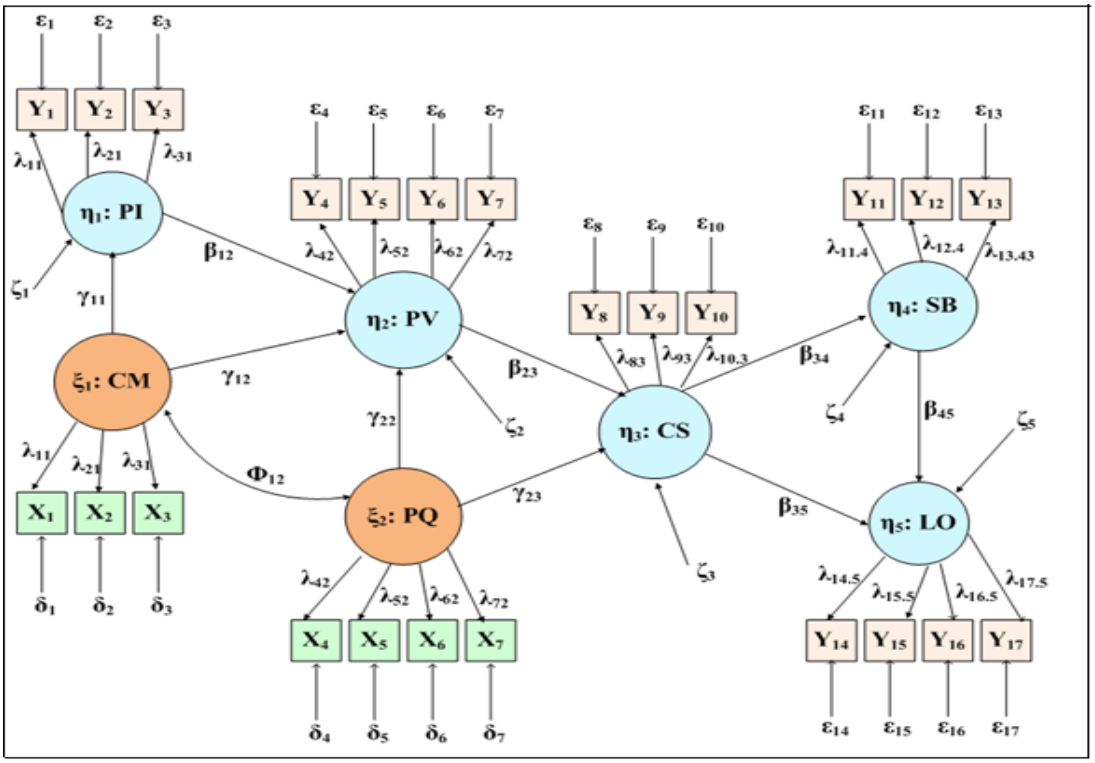


Figure2. SEM of the proposed model

Our survey instrument was a questionnaire with 24 questions, and we conducted a pilot survey on 25 customers for measuring its reliability. According to this survey, the Cronbach's alpha was above 0.7 for all the factors which is an acceptable level of discriminant validity for a survey instrument [29]. The customers chosen and being asked to fill our questionnaire were purchasing goods in large shopping malls, or were waiting for receiving services in a number of service centers[1]. Those who were shopping were asked to specify a product which they used widely. It did not matter if that product was bought during that specific shopping or not, or if the product was made in Iran or another country. The benefit of this part was that a wide range of products were covered by this study and this coverage is of remarkable importance for our study. It should be noted that we had to cover different types of service, because some of them do not cost for the customers and not all aspects of the service could be assessed in such service centers. Therefore, service providers such as post service and restaurant were also included in our study so that the cost of service and its perceived value could also be assessed.

Among the shopping customers, 18% specified facilities and computer hardware, 39% specified grocery, 23% specified health care goods and detergents, and 20% specified clothes and garments. Among the service users, 24% belonged to Bank customers, 15% to travel agencies, 20% to post service users, 25% to Health clinic, and the rest were restaurant customers. Note that equal numbers of questionnaires were distributed among all these service providers, but some of the returned questionnaires were not complete enough and some others were not reliable to be eligible of being included in the analysis, so unequal portions were obtained. As mentioned earlier, since the purpose of this study is to examine if the developed model is applicable for most of the businesses, the diversification of respondents to our survey was of remarkable importance. In other words, the respondents were carefully selected so that main claim-applicability of our model to various types of businesses can be comprehensively examined.

B. Measurement variables

Appendix A shows the questionnaire used for this study which contains measurement variables for each factor of our proposed model. As it can be observed, each of the factors has been measured through a multi-item scale. The explanations of measurement variables are as follows:

There are four items, measuring the perceived quality. The first item measures consistency of the product/service with customized needs of the customer. The second variable measures satisfaction of the customer with the product reliability, lack of defection, and its conformation to the standards. The third item measures expected quality of the product or service, and finally the fourth item measures satisfaction with particular services that are offered by the supplier for using that product/service. (This item tends to assess satisfaction with a supplier compared to other suppliers that he/she may know).

[1] These centers were four Banks; two from public sector and two from private sector, one medical clinic, four fast food restaurants, two travel agencies, and a post service center.

For complaints management, three items are used which measure the customers' opinion about the following cases: first, if the supplier cares about discovering dissatisfaction and its possibly dissatisfied customers. Second if the supplier feels responsive about handling complaints and third if the supplier appreciates reflection of any possible dissatisfaction from all its customers.

The perceived image includes three items: customers' judgment about reputation and prestige of the supplier, his feeling and sense about existence of customer-oriented culture among supplier's personnel, and his feeling about trustworthiness of the supplier.

Four items are used to measure the perceived value: first, customers' judgment about fairness of the price, second, consistency of the price to the quality offered, third the product superiority in meeting specific needs and expectations of the customer, and finally worthiness of the relation with that supplier.

For measuring satisfaction, three items are used: satisfaction with ability of the product in meeting his personal needs and expectations, similarity of that product/service to the ideal product/service of the customer, and overall satisfaction of the customer based on all of his/her considerations.

For measuring the impact of switching barriers, three items are used: First, customer's feeling about the value of his current interactions with the supplier, second his lack of tendency to switch for a new supplier unknown to him, and finally, his unwillingness to switch due to costs and required time for finding a new supplier.

Finally, loyalty was measured through four items: First, repurchasing intention of the customer (his willingness to refer to the same supplier again), his price tolerance (his willingness to refer to the same supplier even with a slight increase in the prices), his tendency for using other products or services offered by the same supplier, and finally, his willingness to introduce and recommend the product or service to others.

C. Structural model analysis and results

We applied LISREL 8.3 for analyzing items of our proposed model and the Maximum Likelihood (ML) as the estimation method. Before any further analysis, we conducted a validity and reliability test in order to check the construct validity of our model.

Convergent validity is obtained in case of higher factor loadings or through Bentler–Bonett normed fit index (BBNFI). This index is defined as the ratio of the difference between chi-square of a given model minus chi-square of a null model divided by model chi-square for the null model[2] [30]. Note that the formula suggested in [31] was used for measuring reliability. Table 1, shows the correlations of the factors, Table 2 presents direct effects (factor loadings for direct relations in the model), Table 3 shows indirect effects along with t-Test values and standard errors for paths of the model, and Table 4 presents the convergent validity and reliability of model components.

Model fit measurement

The goodness-of-fit indices for our proposed model are presented in Table5. These results confirm the consistency of the model assumptions to current set of data. The standard errors, t-test values, and the coefficient of determination (R^2) are presented in Table6. The factor loadings in Table 1 are all statistically significant which implies that all these factors affect satisfaction and loyalty significantly (i.e., the higher the factor loadings are, the stronger that factor will affect satisfaction). Table 6 represents the standard errors, t-Test values and R^2 for model components.

We also used another measurement called the total multivariate coefficient of determination (R^2_m) which is an indicator of the variance of dependent variables (Satisfaction and loyalty) explained by independent variables (other factors of the model which are precedents of satisfaction). This parameter has a value of 0.92 for our proposed model which implies that the independent set of factors explains 92% of the variation in satisfaction and loyalty.

Table1: correlations between factor pairs

	PQ	**CM**	**PI**	**PV**	**CS**	**SB**	**LO**
PQ	1.00						
CM	0.71	1.00					
PI	0.42	0.46	1.00				
PV	0.31	0.35	0.49	1.00			
CS	0.49	0.61	0.52	0.58	1.00		
SB	0.62	0.53	0.64	0.58	0.69	1.00	
LO	0.63	0.47	0.51	0.45	0.76	0.47	1.00

Table2: Direct effects

Direct effects	Parameter Estimates	S.E.	t-Test	R^2
CM-PI	0.786	0.042	5.311	0.845
PI-PV	0.811	0.047	4.229	0.698
CM-PV	0.721	0.038	6.178	0.721
PQ- PV	0.745	0.065	5.328	0.823
PV-CS	0.868	0.039	3.983	0.893
PQ-CS	0.880	0.056	6.456	0.812
CS-SB	0.707	0.021	4.223	0.783
CS-LO	0.794	0.031	4.871	0.814
SB-LO+	0.861	0.025	5.344	0.893
SB-LO-	-0.571	0.032	4.876	0.623

[2] A null model is also called independence model in which correlations are all zero.

Table3: Indirect effects

Indirect causal effects	Parameter Estimates	S.E.	t-Test	R^2
PI-PQ	0. 56	0.045	7.823	0.789
PI-CS	0.72	0.025	6.783	0.677
PI-SB	0.67	0.056	7.213	0.982
PI-LO	0.69	0.032	4.344	0.912
CM-PQ	0.65	0.054	4.767	0.871
CM-SB	-	-	-	-
CM-CS	0.72	0.047	6.145	0.520
CM-LO	0.70	0.049	6.564	0.823
PQ-SB	-	-	-	-
PQ-LO	0.79	0.069	7.347	0.712
PV-SB	-	-	-	-
PV-LO	0.75	0.026	4.257	0.834

Parameter estimations are the basis for concluding about strength of causal relationships among the model factors. As table 2 presents, except for one path (negative switching barriers to loyalty), other factor estimate values are greater than 0.7 which implies strong causal relationships and confirms our hypotheses about relations among the model factors. Besides, the acceptable values obtained for indirect effects in Table 3 (most of which, are above 0.6) also confirm indirect effect of the proposed precedent factors on satisfaction and loyalty.

Model hypotheses testing and results

Based on the factor loadings, model fit measurements, t-test, and p-values, we can conclude that our hypotheses cannot be rejected. The particular hypothesis that we built is the positive relation between satisfaction and switching barriers, which is also confirmed by the results. This result is of remarkable interest in the sense that it provides the following conclusion: if a customer is satisfied enough, satisfaction can perform as a barrier against his/her switching. The reverse of such hypothesis does not necessarily hold because a barrier if considered generally, regardless of being positive or negative, might result in dissatisfaction of the customer. Existence of a barrier, even if it is a strong relationship with the supplier, which is considered as a positive barrier per se, may result in preserving the connection with that supplier without actual satisfaction with the product. A positive barrier is considered as a valuable asset to the customer in the sense that he/she feels happy and safe about continuing that relationship and prefers to keep it for the sake of those barriers.

Table 4: Reliability, discriminant validity and Convergent validity of the model

Factors and variables	Factor loadings	BBNFI	Factor reliability	Alpha (Discriminant Validity)
perceived quality	0.77	0.936	77a	0.78
• meeting customized needs	0.64		0.82	
• conformation to standards expected quality	0.78		0.84	
• conformation to expected quality	0.88		0.70	
• special offers by supplier	0.74		0.76	
complaints management	0.81	0.971	64	0.76
• supplier fondness for finding dissatisfaction	0.61		0.59	
• supplier responsiveness for handling complaint	0.87		0.84	
• supplier appreciation for reflecting dissatisfaction	0.78		0.79	
perceived image	0.79	0.981	69	0.84
• supplier reputation	0.80		0.8	
• customer-oriented culture	0.74		0.61	
• supplier trustworthiness	0.85		0.87	
perceived value	0.80	0.974	73	0.72
• price fairness	0.89		0.83	
• consistency of the quality to price	0.80		0.79	
• Satisfying particular expectations	0.53		0.62	
• worthiness of the relationship with the supplier	0.61		0.61	
Satisfaction	0.89	0.976	87	0.75
• product ability in meeting customized needs	0.70		0.8	
• similarity to customer's ideal product	0.54		0.57	
• overall satisfaction	0.84		0.81	
Switching barriers	0.86	0.979	62	0.78
• worthiness of current interaction	0.87		0.84	
• unwillingness to choose a new supplier	0.71		0.74	
• expensiveness of finding new supplier	0.70		0.68	
Loyalty	0.87	0.982	82	0.81
• willingness to repurchase	0.83		0.79	
• price tolerance	0.57		0.63	
• Willingness to use other products of the same supplier	0.63		0.69	
• Willingness to recommend the supplier to other customers	0.88		0.74	

Table5: The fitness indices for the proposed model

Chi-Square	d.f.	X^2/d.f.	p-Value	CFI	RMSEA	BBNFI	BBNNFI
33.831	15	2.254	< 0.01	0.921	0.043	0.921	0.916

Table6: The standard errors, t-Test values and determination coefficients for model factors

Factors and variables	S.E.	t-Test	R^2
Item-factor effects			
perceived quality	0.031	8.221	
• meeting customized needs	0.011		0.722
• conformation to standards expected quality	0.020		0.793
• conformation to expected quality	0.023		0.767
• special offers by supplier	0.034		0.891
complaints management	0.051	7.118	
• supplier fondness for finding dissatisfaction	0.036		0.723
• supplier responsiveness for handling complaint	0.043		0.618
• supplier appreciation for reflecting dissatisfaction	0.051		0.681
perceived image	0.024	7.211	
• supplier reputation	0.056		0.798
• customer-oriented culture	0.051		0.639
• supplier trustworthiness	0.023		0.756
perceived value	0.087	9.151	
• price fairness	0.034		0.664
• consistency of the quality to price	0.021		0.778
• satisfying particular expectations	0.053		0.512
• worthiness of the relationship with the supplier	0.043		0.611
Satisfaction	0.062	8.218	
• product ability in meeting customized needs	0.034		0.827
• similarity to customer's ideal product	0.056		0.617
• overall satisfaction	0.044		0.897
Switching barriers	0.041	7.456	
• worthiness of current interaction	0.058		0.781
• unwillingness to choose a new supplier	0.054		0.873
• expensiveness of finding new supplier	0.030		0.692
Loyalty	0.072	6.522	
• willingness to repurchase	0.021		0.718
• price tolerance	0.065		0.824
• willingness to use other products of the same supplier	0.033		0.892
• willingness to recommend the supplier to other customers	0.050		0.784

On the other hand, a negative barrier merely precludes the customer to leave and the connection continues without his willingness to afford the product or service from the same supplier anymore. If the barrier is a positive one, it will result in satisfaction and loyalty both (as confirmed by our analyses), but if it is a negative one, satisfaction will be decreased while the customer might be considered loyal. The causal relationship between the positive switching barriers and loyalty, confirms that with such barriers the connection will most probably last longer. But, this fact at the same time means that continuing a relation (i.e. loyalty to supplier) is not always a sequence of high satisfaction and likely has only lasted due to some barriers which may not be as effective as the supplier may expect, in long term transactions with its customers.

Predictive relations testing and results

The Factor loadings (parameter estimate values) in Table 4 show the lambda coefficient values of our model (In figure2, the coefficients are presented). These values are the basis for concluding about if a measurement variable is predicting the corresponding factor well enough to be considered as a good and strong predictor for that factor.

The values obtained through analyzing our model imply that similarity of a product to ideal product of the customer (Y_{10}) is not a very good predictor of the corresponding factor (for satisfaction). The other poor predictor is price tolerance for loyalty (Y_{15}), which implies that a loyal customer is not willing that much to pay more for the same product. Of course, since the amount of

increase in price was not mentioned in the questionnaire, this low fit value may have been resulted from misunderstanding or misinterpretation of the underlying question by the respondents. However, the price tolerance is used as one of the loyalty predictors according to past research. Except these two variables, other measurement variables of our model properly predict their relevant factor.

Finally, other particular hypothesis of our study was about indirect effect of expectations on satisfaction. As we mentioned before, expectations were measured through the assumption that they affect perceived quality. The predictive strength of measurement variable (X_3) confirms this assumption and implies that when expectations are met, the perceived quality gets improved and it results in higher levels of satisfaction in turn.

IV. CONCLUSION

In this paper, a multi-dimensional structural model was developed for testing the general factors affecting satisfaction and loyalty in Iran markets. This model assumes that regardless of the product type or the particular service offered, a number of factors matter to the customer that if they are met, the customers' connection to the supplier is more likely to continue. These factors are perceived quality, perceived value, perceived image, complaint management, and switching barriers. Besides, we supposed that satisfaction affects switching barriers and loyalty positively.

Particular hypotheses were built about the causal relations among these factors, and then tested by LISREL software application. Since these factors are latent variables which cannot be directly measured, multi-scale items were developed for measuring each factor. These items are indicator variables which can be explicitly measured and their strength for predicting the underlying factors were tested in this study as well. The analysis results based on our empirical study for customers of Iranian markets confirm the following items:

(a) Except the negative switching barriers, all of the factors in this model positively affect satisfaction and loyalty, which means that by increasing the level of each factor, the level of satisfaction is expected to increase and this higher satisfaction will result in more loyalty of the customer to the current supplier or service provider.

(b) While existence of switching barriers can result in more loyalty and bond of the customer to that supplier, it can negatively affect customers' loyalty. Based on our analysis results, there should be a separation between positive barriers and negative ones. According to this result, the companies willing to have more loyal customers must care about the proper solutions that can result in achievement of loyalty. In other words, though some barriers such as special offers to their known customers can be good promotions for keeping the customers more loyal, but the supplier should be aware of other negative barriers that may be either created by the company or by external parameters. An example of such barriers is the cost of finding a new supplier which even though not created by the current supplier, but can affect loyalty in future; as soon as finding another supplier in any possible way that might or might not accompany any cost to the customer, he/she will switch to the other supplier if not convinced and satisfied enough to continue the current relationship. A solution to prevent such situation is to care about other factors rather than to merely build barriers. Therefore, switching barriers can only be considered as short term remedies for keeping a customer, and a fundamental assessment is required for manufacturers or service providers in order to remarkably and effectively control satisfaction and loyalty. The results of testing prediction strength of our proposed measurement variables reveal that except very few variables, the others properly predict the factor linked to them and therefore can be used as predictors of the same factors in future studies.

(c) Although increased satisfaction positively affects loyalty (both directly, and indirectly through creating switching barriers), there are exceptions that imply the reverse of this fact may not be true. In other words, high level of loyalty may not imply high level of satisfaction and one of the reasons for this fact can be existence of negative switching barriers.

(d) In our model, the results imply that perceived value and perceived quality do not correlate with switching barriers, while perceived image correlates with it. This correlation can be interpreted as the implicit impact of a good image on precluding customer from leaving.

(e) The variable "meeting customized needs of the customer" not only can predict the perceived quality, but also predicts perceived value and this fact can be interpreted as follows: a product/service which satisfies particular personal needs of the customer creates value for him/her both directly and indirectly (by improving his perception about quality of the product/service).

In summary, the precedent factors of our proposed model explain 92% of the variation in the antecedent factors and this high percentage confirms correctness of the hypotheses made about causal relationships among these factors based on our collected data.

V. LIMITATIONS AND FUTURE RESEARCH

A number of limitations accompanied this study, necessitating caution in interpretation of the results. First, although in the sampling stage several product types and services were covered by the questionnaire respondents, however further study and re-sampling in wider communities can be helpful in confirmation of the results. Besides, while the hypotheses are generally confirmed to exist for both products and services, more systematic research can lead to obtaining more assured models for both products and services, and can reveal existence of possible differences between antecedents of satisfaction for

products and services separately, and also can provide the possibility of identifying the differences between requirements of satisfaction for particular products/services. In addition, this study is done in Iranian market environment and though the analysis confirms our proposed hypotheses, precise and customized study is required for other markets, or for Iranian markets under different market conditions. However, empirical study is merely possible by collecting data for validating this model and this limitation lights the possibility of new results if new data are used. An extension of this study can be to integrate more factors and to get help and feedback directly from the customers who are the eventual targets of such studies in order to discover possibly unknown parameters that positively or negatively affect their perceptions, judgments, and behaviors against any supplier.

APPENDIX A. THE MEASUREMENT VARIABLES FOR MODEL FACTORS

1. Perceived Quality
X_1. The product/service meets my specific needs
X_2. The product/service conforms to standards
X_3. The product/service is consistent to what I expected before use
X_4. This supplier/service provider offers more satisfactory services to me in comparison to other suppliers/service providers

2. Complaint Management
X_5. The supplier is interested in discovering if a customer is not satisfied with the product/service
X_6. The supplier feels responsive to quickly handle the complaints and dissatisfaction
X_7. The supplier treats complaining customers very properly and appreciates reflection of the complaints directly to him/her

3. Perceived Image
Y_1. The supplier company is of reputation from my point of view
Y_2. I have noticed that the supplier respects customers and there is a customer-oriented culture among the personnel
Y_3. I feel the supplier is trustworthy and honest to customers

4. Perceived Value
Y_4. I believe that the price is fair
Y_5. I believe that this price fits the quality and other features of the product/service
Y_6. This is a valuable product because it meets my particular expectations and needs
Y_7. I believe that it's worth to keep buying the product/service from this supplier again

5. Customer Satisfaction
Y_8. Generally, I can say that I am satisfied by this product/service
Y_9. I am satisfied with this product/service because it meets most of my specific needs and expectations
Y_{10}. I am satisfied with this product/service because it is very similar to my ideal product

6. Switching Barriers
Y_{11}. I believe that my current connection to the supplier is worth keeping it
Y_{12}. I do not like to change my supplier right now because we know each other for a while
Y_{13}. I do not like to change my supplier right now because it is time/money consuming to try to find another supplier

7. Loyalty
Y_{14}. I may purchase the same product from the same supplier again
Y_{15}. Even if I have to pay a little more for this product/service, I prefer to buy it from this supplier instead of finding a new supplier
Y_{16}. I guess I will consider using more products/services of this supplier in future
Y_{17}. I like to recommend this supplier to other people too

REFERENCES

[1] Churchill, G.A. and Surprenant, C., "An investigation into the determinants of customer satisfaction", *Journal of Marketing Research, Vol. XIX, November*, pp. 491-504, 1982.

[2] Zeithaml, V.A., Berry, L.L. and Parasuraman, A., "The behavioral consequences of service quality", *Journal of Marketing, Vol. 60, April*, pp. 31-46, 1996.

[3] Zeithaml, V. A., "How consumers evaluation processes differ between goods and services", in Donnelly, J.H. and George, W. (Eds), *Marketing of Services*, American Marketing Association, Chicago, IL, pp. 186-90, 1981.

[4] Gilbert, G. R. and Veloutsou, C., "A cross-industry comparison of customer satisfaction", *Journal of Services Marketing, Vol. 20 No. 5*, 298–308, 2006.

[5] Yi, Y., "A critical review of consumer satisfaction", in Zeithaml, V.A. (Ed.), Review of Marketing 1990, American Marketing Association, Chicago, IL, pp. 68-123, 1990.

[6] Fornell, C., "A national customer satisfaction barometer: the Swedish experience", *Journal of Marketing, Vol. 56*, pp. 6-21, 1992.

[7] Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J. and Bryant, B.E., "The American customer satisfaction index: nature, purpose, and findings", *Journal of Marketing, Vol. 60, October*, pp. 7-18, 1996.

[8] Chan, L. K., Hui, Y. V., Lo, H. P., Tse, S. K., Tso, G. K. F. and Wu,M. L., "Consumer satisfaction index: new practice and findings", *European Journal of Marketing, Vol. 37* No. 5/6, 2003.

[9] Kuusik, A. and Varblane, U., "How to avoid customers leaving: the case of the Estonian telecommunication industry", *Baltic Journal of Management, Vol. 4* No. 1, , pp. 66-79, 2009.

[10] Hsu, S.-H., "Developing an index for online customer satisfaction: Adaptation of American Customer Satisfaction Index", Expert Systems with Applications Vol. 34, pp. 3033–3042, 2008.

[11] Rust, R.T. and Zahorik, A.J., "Customer satisfaction, customer retention, and market share", *Journal of Retailing, Vol. 69*, 145–156, 1993.

[12] Bou-Llusar, J.C., Escrig-Tena, A.B., Roca-Puig, V., Beltr´an-Mart´ın, I., "An empirical assessment of the EFQM excellence model: evaluation as a TQM framework relative to the MBNQA model", *Journal of Operations Management*, doi:10.1016/j.jom.2008.04.001

[13] Bollen, k. A., "Structural Equations with Latent Variables", John Wiley & Sons, New York, NY, 1989.

[14] Sohn, S. Y., & Moon, T. H., "Structural equation model for technology commercialization success index", *Technology Forecasting and Social Change, Vol. 70 No. 9*, 885–899, 2003.

[15] Wilson, A., "Attitudes towards customer satisfaction measurement in the retail sector", International Journal of Market Research, Vol. 44 *No. 2*, pp. 213-22, 2002.

[16] Johnson, M.D., Gustafsson, A., Andressen, T.W., Lervik, L. and Cha, J., "The evolution and future of national customer satisfaction index models", *Journal of Economic Psychology, Vol. 22* No. 2, pp. 217-45, 2001.

[17] Martensen, A., Gronholdt, L., & Kristensen, K., "The drivers of customer satisfaction and loyalty: Cross-industry findings from Denmark", *Total Quality Management,Vol. 11*, 544–553, 2000.

[18] Rust, R.T., Inman, J.J., Jia, J. and Zahorik, A., "What you don't know about customer perceived quality: the role of customer expectation distributions", *Market Science, Vol. 18 No. 1*, pp. 77-92, 1999.

[19] Anderson, E.W. and Sullivan, M.W., "The antecedents and consequences of customer satisfaction for firms", *Marketing Science, Vol. 12, spring*, pp. 125-43, 1993.

[20] Zeithaml, V.A. (1988), "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence", *Journal of Marketing, Vol. 52, July*, pp. 2-22.

[21] Parasuraman, A., Zeithaml, V.A. and Berry, L.L., "A multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing, Vol. 64* No. 1, pp. 12-40, 1988.

[22] Tse, D.K. and Wilton, P.C., "Models of consumer satisfaction formation: an extension", *Journal of Marketing Research, Vol. 25*, May, pp. 204-12, 1988.

[23] Patterson, P. G., "A contingency model of behavioral intentions in a services context", *European Journal of Marketing, Vol. 38 No. 9/10*, pp. 1304-1315, 2004.

[24] Swan, J.E. and Oliver, R.L., "Automobile buyer satisfaction with the salesperson related to equity and disconfirmation", in Hunt, H.K. and Day, R.L. (Eds), Consumer Satisfaction, Disconfirmation and Complaining Behavior, Indiana University Press, Bloomington, IN, 1985.

[25] Bearden, W.O. and Teel, J.E., "Selected determinants of consumer satisfaction and complaint reports", *Journal of Marketing Research, Vol. 20, February*, pp. 21-8, 1983.

[26] Oliver, R.L. and Linda, G., "Effects of satisfaction and its antecedents on consumer preference and intention", in Monroe, K.B. (Ed.), *Advances in Consumer Research, Vol. 8*, Association for Consumer Research, Ann Arbor, MI, pp. 88-93, 1988.

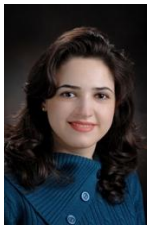
[27] Oliver, R, "*Satisfaction: A behavioral perspective on the consumer"*, New York: McGraw-Hill, 1997.

[28] Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W., *multivariate Data Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1998.

[29] Cronbach, L.J., "Cronbach, Coefficient alpha and the internal structure of tests", Psychometrika 16 (1951), pp. 297–334, 1951.

[30] Bentler, P.M. and Bonett, D.G., "Significance tests and goodness of fit in the analysis of covariance structures", *Psychological Bulletin **88*** , pp. 588–606, 1980.

[31] Fornell, C., Larcker, D.F., "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research 18* (1), 39–50, 1981



Zhila Pirmoradi received the B.S. and M.Sc. degrees in Industrial Engineering from Isfahan University of technology and university of Tehran, in 2006 and 2009 respectively. She started Ph.D. at 2010 and currently is student at Mechatronic Systems Engineering department of Simon Fraser University, Vancouver, Canada.

Her research interests include Customer Satisfaction Management, efficiency improvement of Health Care Systems, and simulation-based optimization.

Monire Jalili received the B.S. and M.Sc. degrees in Industrial Engineering from Khaje Nasir Toosi University in 2006 and University of Tehran in 2009. She will start her PhD at Fall 2010 at University of Central Florida, United States.

Her research interests mainly include Business excellence, Strategic Quality management and Organizational performance management.

Performance Analysis of Routing Protocols for Wireless Ad-Hoc Networks

Reza Qarehbaghi
Simon Fraser University
Vancouver, British Columbia
Canada
E-mail: rqarehba@sfu.ca

***Abstract*— Wireless ad-hoc networks are decentralized wireless networks that do not rely on an infrastructure, such as base stations or access points. Routing protocols in ad-hoc networks specify communication between routers and enable them to select routes between a source and a destination. The choice of the routes is performed by routing algorithms. In this paper, we use OPNET Modeler to simulate three routing protocols for wireless ad-hoc networks in several TCP and UDP scenarios. We analyze route discovery time, end-to-end delay, download response time, and routing traffic overhead in static, less dynamic, and highly dynamic mobility scenarios. Simulation results indicate that the Ad-Hoc On-Demand Distance Vector (AODV) protocol is the most flexible when compared to Dynamic Source Routing (DSR) and Optimized Link State Routing (OLSR) protocols in the case of movement. OLSR is the only protocol that meets the end-to-end delay requirements of less than 20 ms.**

***Index Terms*—AODV, DSR, MANET, OLSR, Wireless ad-hoc network**

I. Introduction

Wireless ad-hoc networks are a collection of mobile nodes that make up a multihop autonomous system. Their decentralized nature makes them suitable for various applications where central node cannot be relied upon. They are made up of multiple nodes or links. Each node requires a route for communication. Hence, each node participates in the routing process by forwarding data to other nodes.

In this paper, we perform a comparative performance analysis of three routing protocols (Ad-Hoc On-Demand Distance Vector (AODV), Dynamic Source Routing (DSR), and Optimized Link State Routing (OLSR)) for mobile ad-hoc networks (MANETs). In Section 2, we survey the literature. Section 3 describes simulated network topology, while simulation scenarios are described in Section 4. Section 5 provides results, and finally, we conclude with Section 6.

II. Ad-hoc Routing Protocols

Ad-hoc routing protocols control routing packets between computing devices in a mobile ad-hoc network. In ad-hoc networks, nodes are not familiar with the topology of the networks and have to discover it [14].

MANET routing protocols can be classified as unicast, multicast, and broadcast. The main goal of unicast protocols is to establish and maintain a route between a pair of nodes. Unicast routing protocols can be further subdivided into reactive (on-demand) and proactive (table-driven) routing protocols based on the method of acquiring information from the other nodes. In addition to these two main groups, there exists hybrid routing protocols that combine the merits of both reactive and proactive routing protocols [1].

Fig. 1 shows a list of ad-hoc routing protocols in each category of proactive, reactive, and hybrid routing protocols.

The advantage of on-demand routing protocols is that they generate less routing overhead compared to table-driven routing protocols. However, a source node may suffer from long delays to obtain a route to a specific destination [1].

The advantage of table-driven routing protocols is that a source node can obtain a routing path immediately if needed. However, they generate a high routing overhead [1].

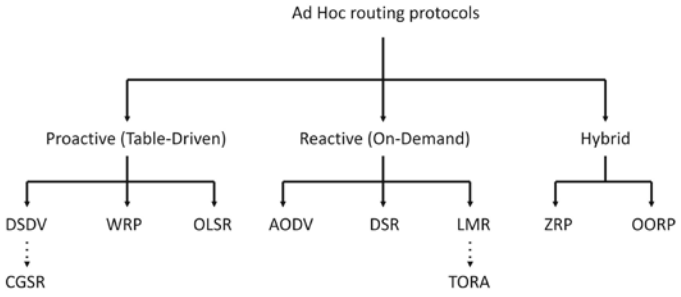


Fig. 1. Ad-hoc routing protocols

A. Ad-Hoc On-Demand Distance Vector (AODV)

AODV is one of the most popular reactive routing protocols. The AODV routing algorithm is suitable for a dynamic self-starting network and users of ad-hoc networks. It ensures loop-free routes even while repairing broken links. Because the protocol does not require global periodic routing advertisements, the overall bandwidth needed for the mobile nodes is considerably smaller than in protocols that need such advertisements [1][6].

AODV defines RREQ, RREP, and RERR message types [1][11]. These message types are received via UDP and, hence, the usual IP header processing applies [11]. When a source node wishes to send a packet to a specific destination, but does not have a valid route to it, it initiates the Path Discovery operation by broadcasting RREQ packet to its neighbors. This request is then forwarded to its neighbors,

until either the destination or an intermediate node with a "fresh enough" route to the destination is located [5]. During the process of forwarding the RREQ message, intermediate nodes record the address of the neighbor from which the first copy of the broadcast packet is received in their route tables. Thus, they establish a reverse path.

After the RREQ reaches the destination or an intermediate node with a fresh enough route, it responds with a route reply (RREP) packet that unicasts the neighbor which first received the RREQ packet, and routes back along the reverse path [5].

When the nodes in the network move and the network topology changes or the links in the active path break, the intermediate node that discovers this link breakage propagates an RERR packet [1][5].

Table I shows the parameters that we used to configure AODV routing protocol in OPNET.

TABLE I
Parameters used for AODV

Attribute	Value	
Route Discovery Parameters	Default	
Active Route Timeout (seconds)	3	
Hello Interval (seconds)	uniform (2, 2.1)	uniform (10, 10.1)
Allowed Hello Loss	2	
Net Diameter	16	

B. *Dynamic Source Routing (DSR)*

Dynamic Source Routing is an on-demand routing protocol based on the concept of source routing where each routed packet carries in its header a complete and ordered list of nodes through which the packet will pass. Hence, intermediate nodes do not need to maintain up-to-date routing information in order to route the packets they forward [3][12].

The protocol consists of two major phases: route discovery and route maintenance [3][12].

When a source node wishes to send a packet to a destination, it obtains a source route by the route discovery mechanism.

At first, a source node consults its route cache to determine whether it already has a route to the destination. If it does not have a route to destination, it initiates route discovery by broadcasting a Route Request (RREQ) packet. The RREQ packet is then answered by a Route Reply (RREP) packet when the RREQ reaches either the destination, or an intermediate node with an unexpired route to the destination in its route cache [3][5][12].

The route maintenance mechanism uses Route Error (RRER) packets and acknowledgments. RERR packets are generated to notify the source node that a source route is broken [3][12].

Table II shows the parameters that we used to configure DSR routing protocol in OPNET.

TABLE II
Parameters used for DSR

Attribute	Value
⊟ Route Cache Parameters	(...)
·· Max Cached Routes	Infinity
·· Route Expiry Timer (seconds)	60 300
⊞ Route Cache Export	Do Not Export
⊞ Send Buffer Parameters	Default
⊟ Route Discovery Parameters	(...)
·· Request Table Size (nodes)	16
·· Maximum Request Table Identifi...	16
·· Maximum Request Retransmissio...	16

C. *Optimized Link State Routing (OSLR)*

Optimized Link State Routing is a proactive routing protocol. It periodically exchanges topology information with other network nodes. The periodic nature of the protocol creates a large amount of overhead, but OLSR reduces this overhead by using "Multi Point Relays" (MPR). Each node selects a set of neighbouring nodes as MPRs and only those MPRs are responsible for forwarding routing messages and network wide traffic. In OLSR, only nodes that have been selected as MPRs by a neighbouring nodes, announce this information periodically. Hence, the network knows the selected node has the ability to reach the nodes that have selected it as an MPR [1][13].

OLSR does not require reliable control message delivery because each node sends control messages periodically and can thus sustain reasonable loss of control messages. Each control message uses a sequence number, which is incremented for each message. Therefore, the protocol does not require sequenced delivery of messages [1][13]. OLSR uses Topology Control (TC) messages to provide sufficient link state information to every network node to allow route calculation [1][13].

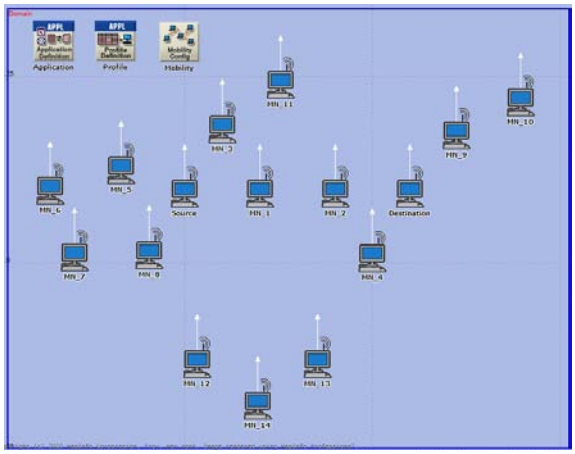
Table III shows the parameters that we used to configure OLSR routing protocol in OPNET.

TABLE III
Parameters used for OLSR

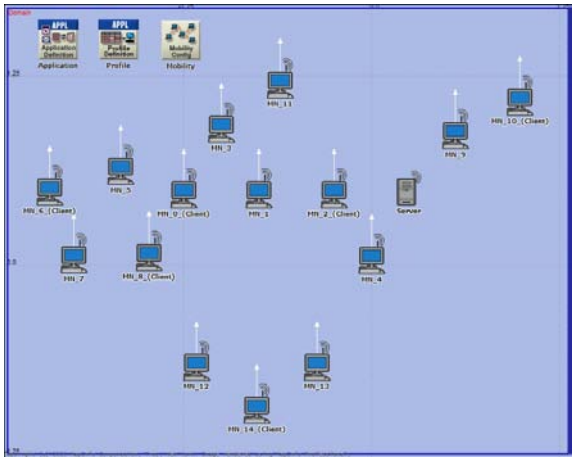
Attribute	Value
Willingness	Willingness Default
Hello Interval (seconds)	2 10 2 10
TC Interval (seconds)	5 5 25 25
Neighbor Hold Time (seconds)	6.0
Topology Hold Time (seconds)	15.0
Duplicate Message Hold Time (seconds)	30.0

III. OPNET Simulated Network Topologies

We developed an ad-hoc network OPNET model to evaluate the performance of three routing protocols in various environments.



(a)



(b)

Fig. 2. Generic OPNET model of: (a) UDP connection scenario (b) TCP connection scenario.

The simulated ad-hoc network shown in Fig. 2(a) is a generic model of a UDP connection scenario for highly dynamic environments. The simulated ad-hoc network shown in Fig. 2(b) is a generic model of a TCP connection scenario for highly dynamic environments. Note that the routing protocol and mobility differ in each scenario. Every scenario consists of 16 wireless LAN (WLAN) nodes connected via a low power wireless connection with a transmission power of 5 mW and a packet reception-power threshold of -90 dbm. Each node covers an area of about 675 m [4], and as the distance between neighboring nodes is about 500 m, each node can only see its neighboring nodes.

A. Node Mobility

Mobility is one of the main factors in ad-hoc networks. Modeling the movement of a set of nodes is essential for evaluating performance of an ad-hoc network. For this project, we used a random waypoint model available in OPNET Modeler.

We created three mobility models: static, less dynamic, and highly dynamic. In the first mobility model, all of the nodes are static, and hence, we can identify the best results for each routing protocol. In the second mobility model, we used a very low-speed movement (e.g, walking speed) for a few mobile nodes. In the third mobility model, we created a high-speed movement (e.g, driving in the city) for all mobile nodes. We identified routing protocol that performs best in different dynamic environments and the routing protocol that is more stable in various mobility scenarios.

IV. Simulation Scenarios

We used the OPNET Modeler 16.0.A to simulate wireless ad-hoc networks with three routing protocols: AODV, DSR, and OLSR. We also analyzed the effect of periodic routing advertisement in different movement conditions.

The first scenario is a static scenario used to analyzing the static ad-hoc network and compares its performance with other scenarios. In the case of the static network, nodes are motionless and thus, all routes are valid for the entire simulation time.

In the second scenario, some nodes move with a very low speed. The speed is comparable to the human walk (about 1 m/s). In this paper, this scenario is called the less dynamic scenario.

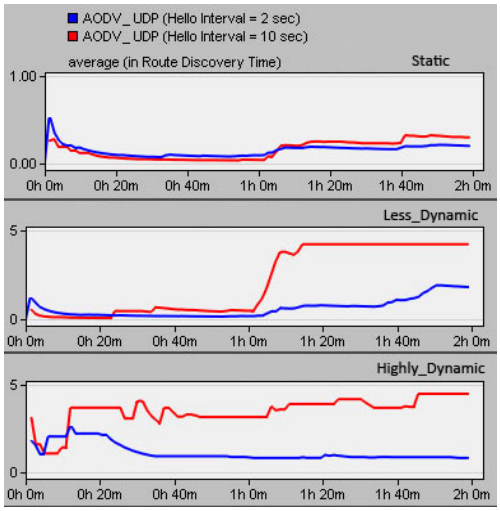
The third scenario includes high-speed nodes that move with maximum speed equal to the speed of the cars in the city (about 50 km/h). The nodes may or may not move during this simulation run.

For each static, less dynamic, and highly dynamic scenarios, we used two types of connections (UDP and TCP) and three ad-hoc routing protocols with various attributes (Tables I, II, and III).

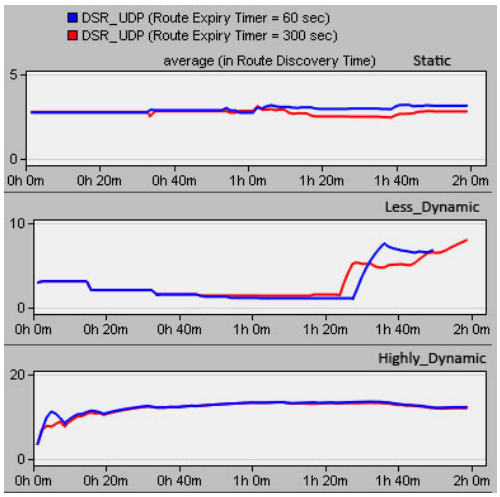
In total, we created 48 simulation scenarios, and analyzed route discovery time, end-to-end delay, download response time, and routing traffic overhead in TCP and UDP connection scenarios.

A. UDP connection OPNET model

In UDP connection scenarios, a two-hour interval of the Matrix III movie trace [10] streamed from the source node to the destination node using various routing protocols and attributes. We created 24 simulation scenarios for the UDP connection case. As it is not possible to compare all 24 scenarios in the paper, we chose only nine scenarios to compare.



(a)



(b)

Fig. 3. Average route discovery time in the UDP connection scenarios for (a) AODV (b) DSR.

To identify the best video streaming performance related to a variety of AODV and DSR attributes, we examined route discovery time because the faster the nodes can locate a route, the faster they can begin sending the video which cause lower end-to-end delay. As shown in Fig. 3(a) and 3(b), each scenario has a different route discovery time. In case of AODV scenarios (Fig. 3.a), the AODV routing protocol with the hello message interval of 2 seconds has better route discovery time as compared to other scenarios. In every DSR scenario (Fig. 3.b), the DSR routing protocol with the route expiry timer of 300 seconds has better route discovery time.

OLSR is a proactive routing protocol, and hence, it has a route to the destination before it begins sending data. This method results in smaller end-to-end delay in streaming video packets in comparison to AODV and DSR.

As shown in Fig. 4, we compared packet end-to-end delay in different scenarios with OLSR routing protocol. The OLSR routing protocol with the hello message interval of 2 seconds and the topology control message interval of 5 seconds performed better in finding a route to the destination and in dealing with the movement.

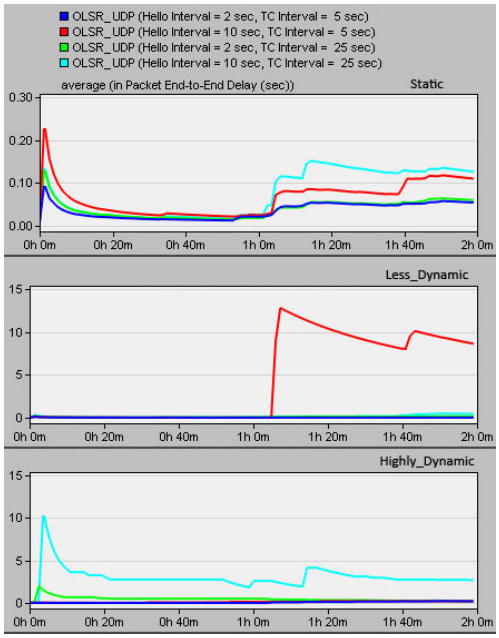


Fig. 4. Average packet end-to-end delay in the UDP connection scenarios for OLSR.

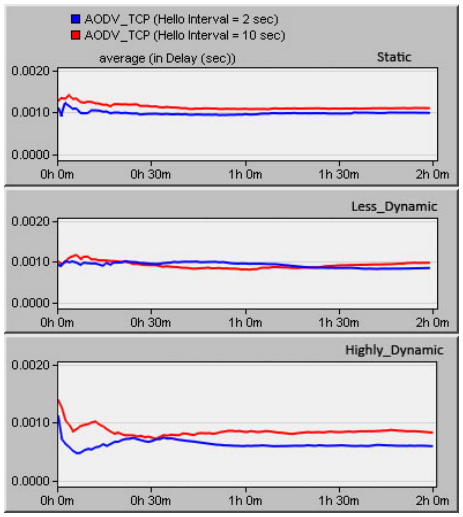
B. TCP connection OPNET model

In the TCP connection scenarios, we used six client nodes that download 50 KB of data from the server with different start times having a uniform distribution of 20 to 80 seconds. After 180 seconds, the client nodes restart downloading the data.

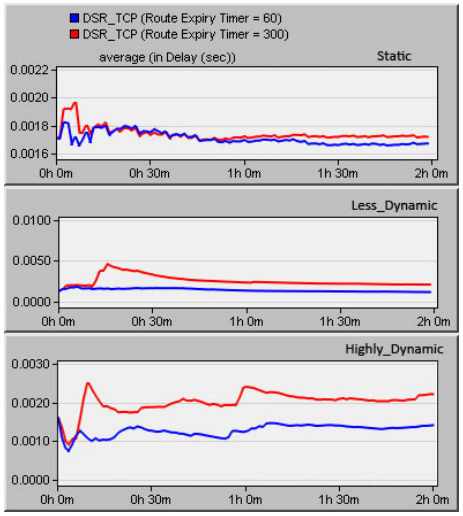
In the scenarios with the DSR routing protocol that we used two route expiry timers of 60 and 300 seconds, In the first case, the route to the FTP server expires after 60 seconds and to send another file, the source node has to locate a new route to it. In the second case, the route to the server stays valid and the client node can use the route to request another file.

For ease of comparision, we chose one scenario in each routing protocol scenario. Here, unlike the UDP connection scenario, we consider wireless delay.

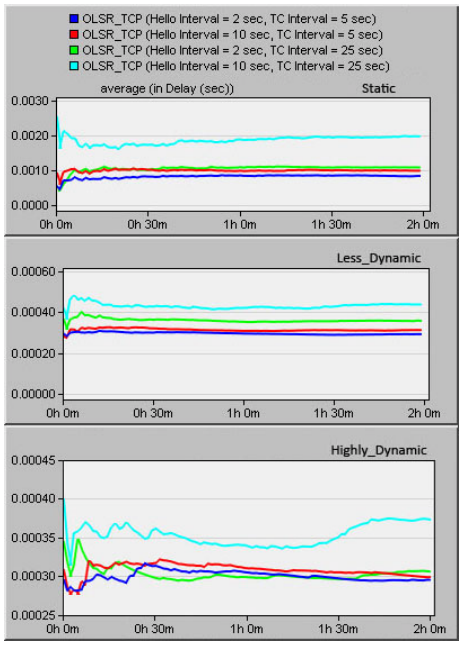
The average wireless delays for AODV, DSR, and OLSR are shown in Fig. 5. The AODV with the hello message interval of 2 seconds has minimum wireless delay, as shown in Fig. 5(a). The DSR with the route expiry timer of 60 seconds has minimum wireless delay, as shown in Fugue 5(b). The OLSR routing protocol with the hello message interval of 2 seconds and the TC message interval of 5 seconds has minimum wireless delay, as shown in Fig. 5(c).



(a)



(b)



(c)

Fig. 5. Average wireless delay in the TCP connection ad-hoc network for (a) AODV (b) DSR (c) OLSR.

V. RESULTS

This section provides some results of our comparison between routing protocols. We first consider route discovery time as one of the most important factors in the on-demand routing protocols. We then consider packet end-to-end delay in the UDP connection scenario which is streaming a video over the network, and download response time in the TCP connection scenario which is downloading a file from FTP server through some of the network nodes. Finally, we show the routing traffic overhead generated throughout the ad-hoc network using each routing protocol.

A. Route Discovery Time

Route discovery time is an important factor in on-demand routing protocols. It causes large delays if the route discovery operation fails to find a route to the destination.

The route discovery time for AODV and DSR routing protocols in all UDP connection scenarios is shown in Fig. 6.

In the static UDP scenario, the route discovery phase in AODV performed approximately 10 times faster than the route discovery phase of DSR in the static ad-hoc network. In the case of the less dynamic UDP network, once again AODV showed less route discovery time, which is similar to the highly dynamic scenario.

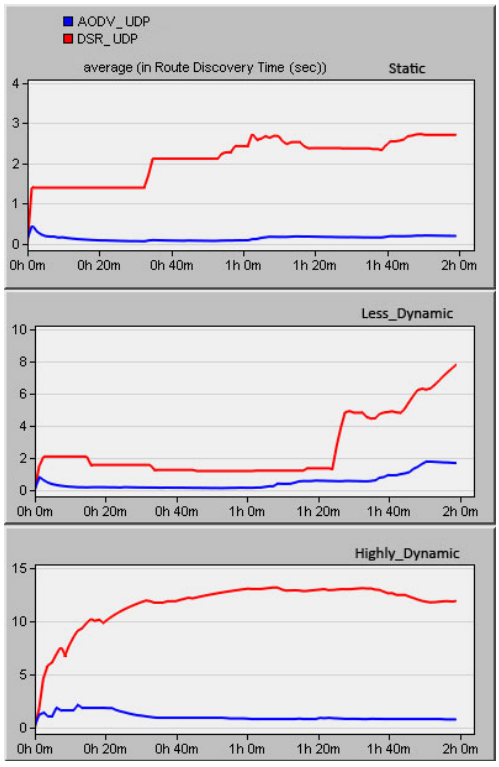


Fig. 6. Average route discovery time (sec) in all UDP connection ad-hoc network scenarios for AODV and DSR.

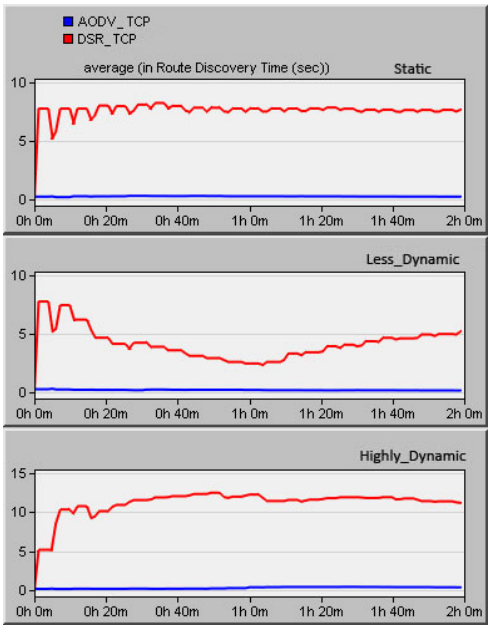


Fig. 7. Average route discovery time (sec) in all TCP connection ad-hoc network scenarios for AODV and DSR.

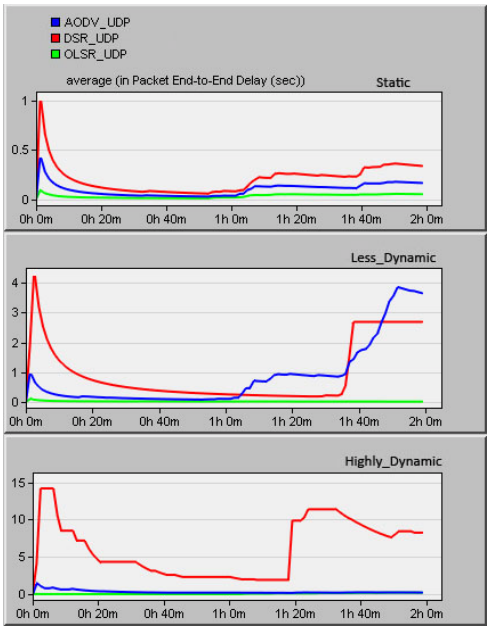


Fig. 8. Average packet end-to-end delay (sec) in all UDP connection ad-hoc network scenarios for AODV, DSR, and OLSR.

The route discovery phase in the AODV routing protocol is independent of the network topology, as shown in Fig. 6. This implies that the changes in network topology do not affect the performance of the route discovery phase of AODV in a video streaming network. Unlike AODV, the route discovery phase in DSR is dependent on the network topology. The route discovery time in DSR is higher in scenarios that include movements and the continuously changing topology, as clearly shown in Fig. 6.

Results shown in Fig. 7 indicate that for all TCP connection scenarios, finding routes for the AODV routing protocol are almost instantaneous, while DSR routing protocol takes more time to discover routes.

B. End-to-end delay/Download response time

Fig. 8 demonstrates the end-to-end delay results for the UDP connection scenarios. The end-to-end delay in the static network for all three routing protocols is approximately less than half a second for most of the simulation time. As we expected from the route discovery time, the AODV end-to-end delay is almost constant for all mobilities. However, the DSR end-to-end delay grows as mobility increases.

The OLSR has the smallest delay in all mobility scenarios because it is a proactive routing protocol and it discovers routes before attempting to send any data. Hence, OLSR routing protocol has a general overview of network topology, and every node in OLSR has at least one valid route to each reachable destination.

The download response time in TCP connection scenarios is shown in Fig. 9. Although the route discovery time for DSR is higher than the AODV route discovery time, it has the smallest download response time. The DSR download response time is also smaller than the download response time in OLSR.

DSR has the smallest download response time, and this fact might be due to the DSR source routing. DSR uses source routing, which implies that a destination node does not need to discover a new route to the source node in order to send back the acknowledgement for each TCP packet, and the path to the source is the reverse path of the source to the destination path.

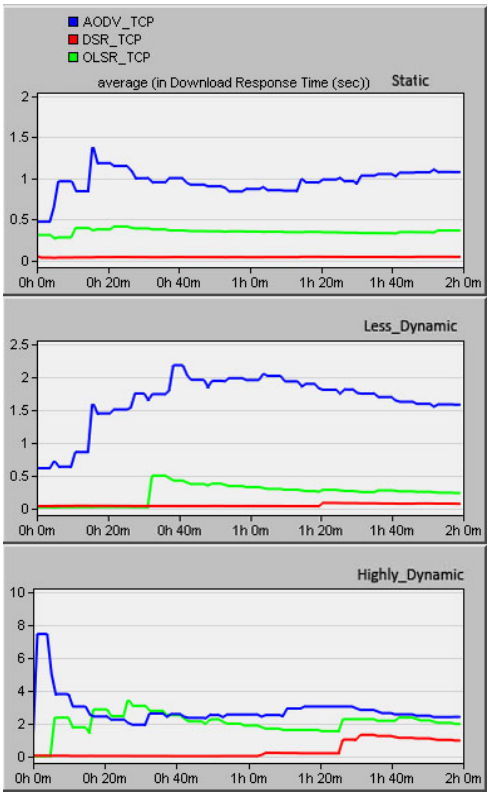


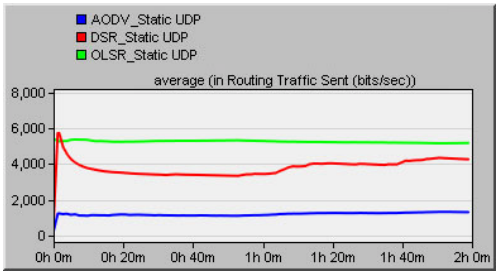
Fig. 9. Average download response time (sec) in all TCP connection ad-hoc network scenarios for AODV, DSR, and OLSR.

C. *Routing Traffic overhead*

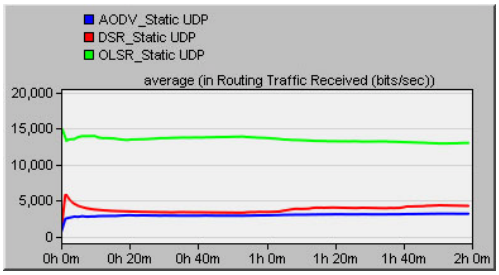
Routing traffic is another important factor in ad-hoc networks, and can be a determinant factor in networks such as Wireless Sensor Networks (WSNs).

Average routing traffic sent and received throughout the UDP connection ad-hoc network for AODV, DSR, and OLSR is shown in Fig. 10. As expected, OLSR has much larger routing traffic overhead in comparison to AODV and DSR.

Fig. 10 also illustrates an interesting result in the OLSR routing protocol. The protocol sends about 5,500 bits/second it receives about 14,000 bits/sec of routing traffic, which implies each packet is duplicated approximately three times in the network. Note that the routing traffic overhead sent and received is constant in DSR (and is approximately the same in AODV).



(a)

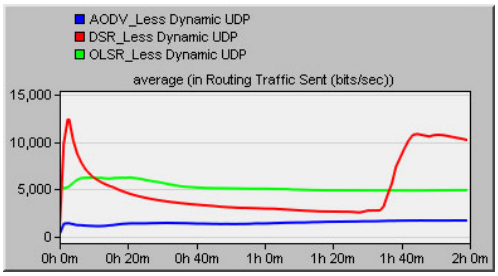


(b)

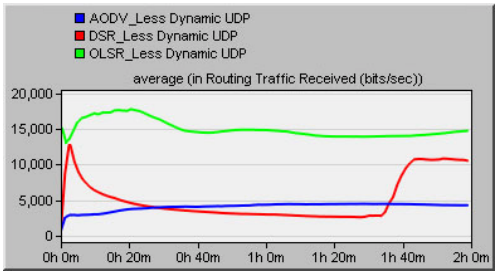
Fig. 10. Average routing traffic sent (a) and received (b) in the static ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.

The results for the less dynamic UDP connection ad-hoc network are shown in Fig. 11. The only difference in the routing traffic sent and received in the less dynamic UDP connection network compared to the static UDP connection network is a slight increase in routing traffic sent and received.

The tendency of DSR to send more routing traffic throughout the network in comparison to the static network is shown in Fig. 11. The DSR sends much more routing traffic in the presence of highly dynamic nodes and in the video streaming scenario, as shown in Fig. 12.

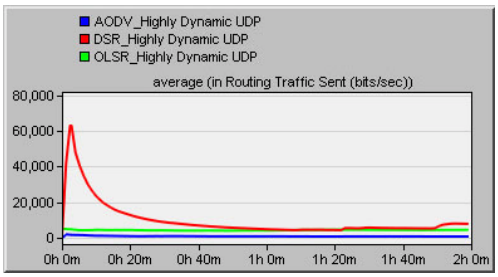


(a)

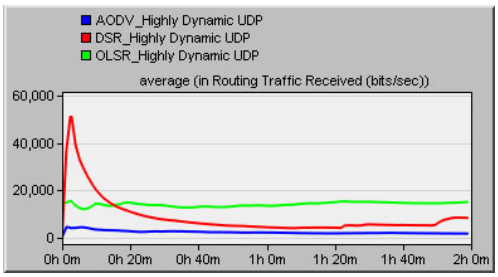


(b)

Fig. 11. Average routing traffic sent (a) and received (b) in the less dynamic ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.



(a)

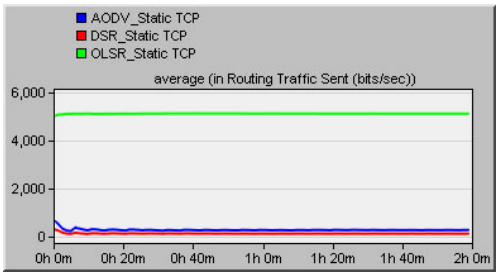


(b)

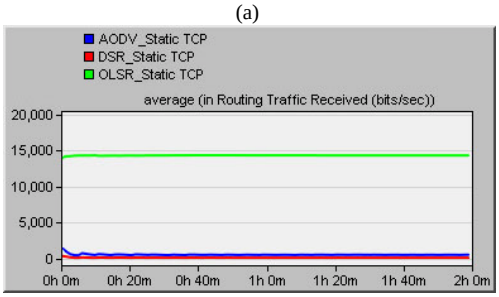
Fig. 12. Average routing traffic sent (a) and received (b) in the highly dynamic ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.

Fig.s 13, Fig. 14, and Fig. 15 show the average routing traffic sent and received throughout the ad-hoc network in the TCP connection scenario. These results show that although DSR routing traffic increasing as node movement increases in the video streaming scenario, DSR has consistent results in the file downloading scenario, and it has the least amount of routing traffic compared to AODV and OLSR.

Furthermore, OLSR shows the same results as the UDP connection scenario and it has a very large amount of traffic sent and received.

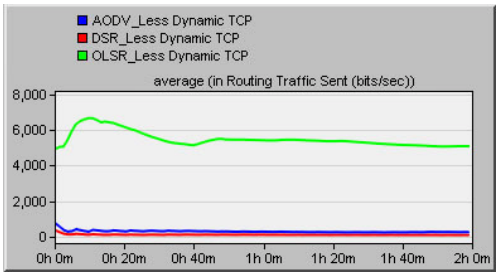


(a)

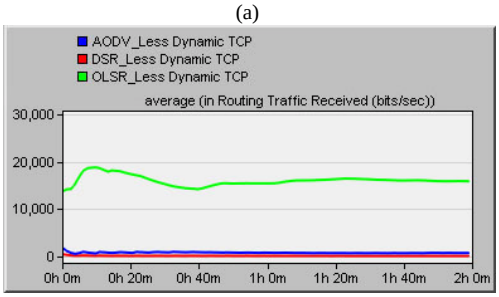


(b)

Fig. 13. Average routing traffic sent (a) and received (b) in the static ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.

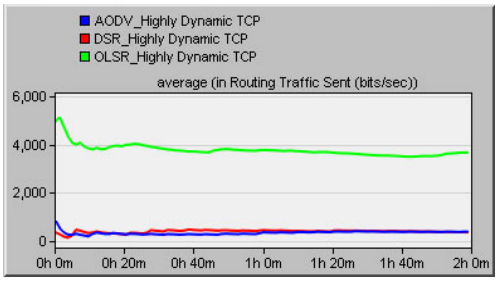


(a)

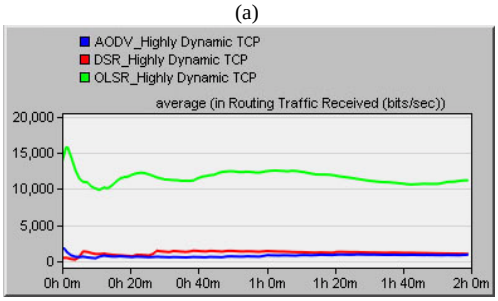


(b)

Fig. 14. Average routing traffic sent (a) and received (b) in the less dynamic ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.



(a)



(b)

Fig. 15. Average routing traffic sent (a) and received (b) in the highly dynamic ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.

VI. Conclusion

In this paper, we simulated and compared the performance of various wireless ad-hoc routing protocols in various environments. Generally, AODV is the most flexible protocol of the protocols that we simulated, and it works better in the presence of movement while generating low routing traffic overhead.

AODV: Simulation results indicate that AODV is the most flexible routing protocol in the presence of movement. It also has the least routing traffic overhead in most of the scenarios, and its routing discovery phase can discover routes very quickly.

DSR: In video streaming scenarios, DSR does not exhibit a good performance, especially in the presence of movement. In addition to large end-to-end delays in video streaming in the DSR routing protocol, DSR also suffers from less flexibility in the presence of movement, and as the movement increases, the route discovery time and routing traffic overhead increase.

In the case of the TCP connection scenarios, DSR demonstrates good performance only in download response time and has low routing traffic overhead, and rather than being inflexible in video streaming scenarios, it can handle movement in the network used for transferring files.

OLSR: In the UDP connection simulation scenario which is streaming a video over the network, only the OLSR routing protocol maintained the demand for an ideal end-to-end delay value that is less than 20 ms [10], while other protocols suffered from large end-to-end delays. In order to accomplish that, OLSR generates a large routing traffic overhead to have always at least one route to every node in the network.

In the case of TCP connection scenarios, OLSR does not demonstrate good performance. Despite generating a large amount of overhead to know always the network topology, it

does not have smaller download response time than DSR.

In the presence of movement, DSR and OLSR impose significant routing traffic overhead in ad-hoc networks, and may fail to perform in conditions with limited resources available (e.g, buffer, bandwidth, power, etc).

REFERENCES

[1] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, Ad Hoc Mobile Wireless Networks: Principles, Protocols, and Applications, New York, Auerbach Publications, 2007.

[2] G. Jayakumar and G. Ganapathy, "Performance Comparison of Mobile Ad-hoc Network Routing Protocol," *IJCSNS International Journal of Computer Science and Network Security*, vol.7, no.11, pp. 77-84, Nov 2007.

[3] J. Broch *et al.*, "A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, Dallas, Texas, United States, October 1998, pp. 85–97.

[4] M. T. Hyland, "Performance Evaluation of Ad Hoc Routing Protocols in a Swarm of Autonomous Unmanned Aerial Vehicles," M.S. Thesis, Air Force Institute of Technology, Ohio, USA, Mar. 2007.

[5] E. Royer and C. Toh, "A Review of Current Routing Protocols for Ad-hoc Mobile Wireless Networks," IEEE Personal Communication Magazine, vol. 6, pp. 46-55, April 1999.

[6] A. Suresh, "Performance Analysis of Ad hoc On-demand Distance Vector routing (AODV) using OPNET Simulator," M.S. Mini Project, University of Bremen, Bremen, Germany, 2005.

[7] K. Gorantala, "Routing Protocols in Mobile Ad-hoc Networks," M.S. Thesis, Umeå University, Sweden, 2006.

[8] Cisco Systems, Inc., *Cisco Network Planning Solution Standard Models User Guide*, Cisco Systems, Inc., 2005.

[9] OPNET Technologies, Inc. Making Networks and Applications Perform, "HOW TO: Design Mobile Ad Hoc Networks and Protocols," OPNET Technologies, Inc., 2007.

[10] W. Hrudey and Lj. Trajkovic, "Streaming video content over IEEE 802.16/WiMAX broadband access," *OPNETWORK 2008*, Washington, DC, Aug. 2008.

[11] Ad hoc On-Demand Distance Vector (AODV) Routing [online]. Available: http://www.ietf.org/rfc/rfc3561.txt.

[12] The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4 [online]. Available: http://www.ietf.org/rfc/rfc4728.txt.

[13] Optimized Link State Routing Protocol (OLSR) [online]. Available: http://www.ietf.org/rfc/rfc3626.txt.

[14] List of ad-hoc routing protocols [online]. Available: http://en.wikipedia.org/wiki/List_of_ad-hoc_routing_protocols.

A Novel Coarse-to-fine Global Motion Estimation In Video

Chun Qian and Ivan Bajic, *Member, IEEE*

***Abstract*—We proposed a three-stage coarse-to-fine approach undergoing translational, affine, and perspective transformation. The coarse-to-fine approach allows calculation of global motion parameters (GMPs) from a lower to higher approximation level in order to reduce the computational complexity. The proposed GME method is also able to deal with motion ambiguity caused by fast camera zooming. Oriented in video background segmentation, our algorithm will achieve the minimum mean error of video background rather than the entire frame. Simulation test results have shown that our proposed GME method outperforms conventional MPEG-4 GME in algorithm efficiency and accuracy. The abstract goes here.**

***Index Terms*—global motion estimation(GME), global motion compensation(GMC), coarse-to-fine, affine, perspective,zoom scale, gradient-descent, outlier.**

I. Introduction

The content-based representation introduced in MPEG-4[1] allows for encoding video frames as a stationary background mosaic plus a collection of moving objects (MOs)[11]. This representation of the video is a true content-based description, and allows object-based manipulation and adaptation, as proposed in MPEG-7[2].

Global motion estimation (GME) was originally introduced in MPEG-4 advanced simple profile (ASP) for improving the coding efficiency[17][18]. Global motion is usually caused by camera movement such as pan, tilt, rotation, and zooming. With global motion estimation, two pictures of the same scene could be related by a non-singular nonlinear transformation of the projective plane. Estimating the camera motion is called global motion estimation (GME), while the transformation parameters are called global motion parameters(GMP).

GME has many applications, such as video object segmentation[11], content-based video retrieval[12], coding applications[8][9][14], and entertainment[10]. One of the most important applications is GME based dynamic background sprite[1]. Considering the situation in Figure 1, a camera moves from $C_0(0,0,0)$ to $C_1(t_x,t_y,t_z)$ in the C_0 coordinate system (X,Y,Z), and follows a rotation. Suppose two pictures $I(x,y)$ and $I'(x',y')$ are taken at $C_0(0,0,0)$ and $C_1(t_x,t_y,t_z)$ of the same scene, and the corresponding image coordinates of point P in *I* and *I'* are (x,y) and (x',y'), respectively. The corresponding transformation can be described as a relationship between the two sets of image coordinates:

$$x' = \frac{r_{11}\frac{f'}{f}x + r_{12}\frac{f'}{f}y + r_{13}f' + \frac{t_X f'}{Z}}{\frac{r_{31}}{f}x + \frac{r_{32}}{f}y + r_{33} + \frac{t_Z}{Z}},$$
$$y' = \frac{r_{21}\frac{f'}{f}x + r_{22}\frac{f'}{f}y + r_{23}f' + \frac{t_Y f'}{Z}}{\frac{r_{31}}{f}x + \frac{r_{32}}{f}y + r_{33} + \frac{t_Z}{Z}}, \quad (1)$$

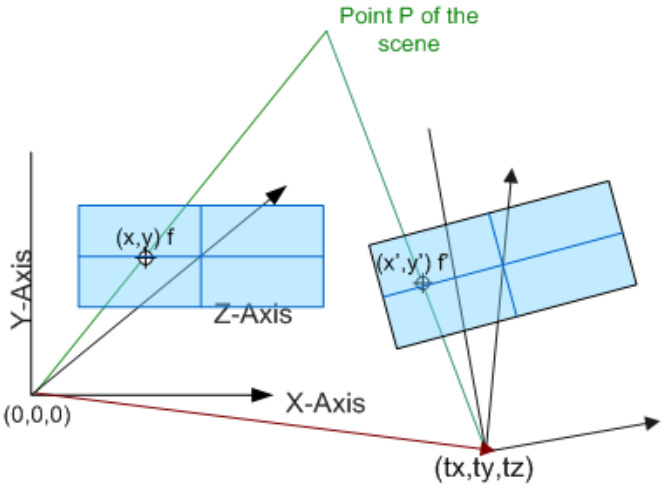


Fig. 1: Principle of Global Motion Estimation

here f and f' are the focal lengths at C_0 and C_1, respectively; and r_{ij} is the element of rotation matrix *R*.

$$R(\varphi_x,\varphi_y,\varphi_z) = \begin{pmatrix} \cos\varphi_z & -\sin\varphi_z & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_x & -\sin\varphi_x \\ 0 & \sin\varphi_x & \cos\varphi_x \end{pmatrix} \times \begin{pmatrix} \cos\varphi_y & 0 & \sin\varphi_y \\ 0 & 1 & 0 \\ sin\varphi_y & 0 & \cos\varphi_y \end{pmatrix}, \quad (2)$$

here $\varphi_x, \varphi_y, \varphi_z$ are the rotation angles around X-, Y-, and Z-axis, respectively.

Equation(1) can be simplified into the following general eight-parameter perspective transformation:

$$x' = \frac{p_1 x + p_2 y + p_3}{p_7 x + p_8 y + 1},$$
$$y' = \frac{p_4 x + p_5 y + p_6}{p_7 x + p_8 y + 1}, \quad (3)$$

or in the following matrix form:

$$\frac{1}{\omega}\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \frac{1}{\omega}\begin{pmatrix} p_1 p_2 p_3 \\ p_4 p_5 p_6 \\ p_7 p_8\ 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad (4)$$

where $\omega = px + py + 1$ is a scaling parameter that describes the non-linear feature of the perspective transformation. From equation(1) and (3), it is clear that parameters $(p_1,...,p_8)$ are cross-related to each other. When (p_7,p_8)=$(0,0)$, the perspective transformation becomes a general affine transformation.

The existing GME methods can be classified into two categories: (1) feature-based approaches and (2) photometric consistency-based approaches[6]. The goal of both approaches is to find the optimal parameter set of a specified transformation model that minimizes some registration error metric. For the feature based method, the metric is usually the distance between the transformed feature points and previously tracked features. For the photometric consistency-based registration, the metric is often based on the normalized correlation measures between the overlapping areas of the compensated images (video frames).

Feature-based methods are very popular due to their low computational complexity[**?**], [**?**], however they heavily rely on the strong correspondences between distinctive feature points in the two images (frames), such as points on the corners and other geometric features. Therefore, feature-based methods could be very powerful in sport field images or other scenes where distinctive geometric features can be expected.

Photometric consistency-based approaches use image registration to find an optimal parameter set of a perspective transformation of the anchor image, so that the correlation between the target image and the compensated anchor images is maximized.

In summary, the photometric approaches are based on the entire frame and they usually integrate some measure of similarity between colors or illumination values. The feature-based methods are based only on certain sparse feature points and are based on spatial distance measures.

In photometric approaches, to maximize normalized cross-correlation between two images (frames), the absolute or squared displaced frame difference (DFD) is often used for its simplicity and efficiency for error energy minimization. A typical cost function E is to be minimized with respect to the transformation parameter vector $\boldsymbol{p}$.

Based on image registration, GME can be described by a minimization of the image registration error:

$$\begin{aligned} E(\boldsymbol{p}, I, I') &= \frac{1}{N_R} \sum |I(x,y) - I'(x',y')| \\ &= \frac{1}{N_R} \sum |I(x,y) - I'(\boldsymbol{T}(\boldsymbol{p}; x, y))|, \end{aligned} \tag{5}$$

where T is the transformation matrix with parameter vector $\boldsymbol{p}$. R is the overlapped area between the two global motion compensated frames.

The minimization of the cost function, E, is a non-linear optimization problem that can be solved by gradient descent methods. Starting with an initial parameter vector p, each gradient descent optimization terminates when the change of error term E becomes less than a pre-defined threshold,

$$|E(\boldsymbol{p}^{n+1}) - E(\boldsymbol{p}^n)| < \varepsilon. \tag{6}$$

In our experiments, the threshold value was set to 0.001.

A. Overview of Conventional MPEG-4 GME

Conventional MPEG-4 GME is a photometric consistency-based approach using gradient descent optimization. Because of large parameter displacement, gradient descent optimization could be trapped in a local minimum when it attempts to find an optimal parameter value far from the starting point. Considering the relative independent of parameter (p_3, p_6), the conventional MPEG-4 GME[19][21] adopts the following two-stage approach to ensure the global convergence, as shown in figure 2.

1) Initial translational stage: Obtain a coarse estimation of (p_3, p_6) by using translational motion search; this will provide a good starting point for the following stage.
2) Perspective stage: Calculate all eight parameters using gradient descent based optimization, and the gradient descent step will be repeated up to a maximum 32 loops till the parameter value updating is smaller than a predefined threshold.

In each stage, a histogram of an global motion compensated error frame, $|e(i,j)|$, is constructed to calculate a threshold,T; this threshold is so chosen that the top 10 percent of pixels having $|e(i,j)| > T$ will be excluded from the following iterations. This outlier rejection step is applied to each gradient descent loop

II. Proposed Three-stage Coarse-to-fine GME

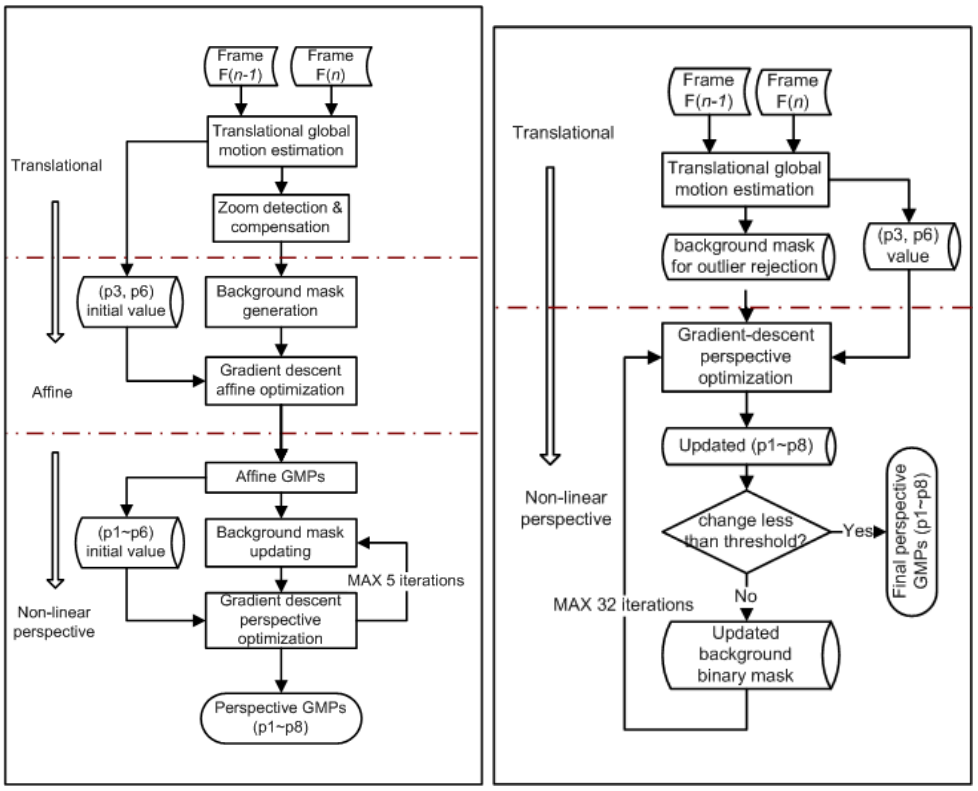
The conventional MPEG-4 GME can achieve a high accuracy, however its computational cost is also very high due to the exhaustive iteration loops; furthermore it is often trapped in local minimum when there is a large scale camera zooming. In order to improve algorithm efficiency and accuracy, we proposed a three-stage approach. The purpose of this three-stage approach is to estimate perspective parameters at different stage by using lower to higher level camera motion approximation model, at the same time due to the cross-correlation among the global perspective parameters. By grouping the perspective GMPs into translational, affine, and perspective group, we can treat each group as a local parameter while providing these values as good initial guesses for the next parameter group.

Considering the slow camera panning(φ_y), tilting(φ_x), and rotation(φ_z) factors between two consecutive video frames, we can obtain the following parameter mapping by using small angle approximation to equation (1).

$$\begin{pmatrix} p_1 p_2 p_3 \\ p_4 p_5 p_6 \\ p_7 p_8\ 1 \end{pmatrix} = \begin{pmatrix} s & s\varphi_z & sf\varphi_y + t_x/Z \\ -s\varphi_z & s & -sf\varphi_x + t_y/Z \\ -\varphi_y/f\varphi_x/f & & 1 \end{pmatrix}, \tag{7}$$

here s=f'/f is a camera zoom scale. From equation(7), it is clear that (p_1, p_2, p_4, p_5) are typical rigid affine parameter(φ_z, s); and (p_3, p_6) and (p_7, p_8) are two relative independent group. Therefore it is possible to group $(p_7, ..., p_8)$ into different motion component group and to calculate them separately. Based on this assumption, our proposed strategy proceeds with the following three main stages:

1) Translational stage: To Estimate of coarse global motion vector (p_3, p_6) using sub-sampled video frame. Then a global motion vector refinement combined with coarse zooming scale detection step can not only align most of the adjacent frames but also provides a good starting point for affine parameters of the next stage.



(a) Conventional MPEG-4 (b) Proposed method

Fig. 2: Global Motion Estimation methods

Fig. 3: Updating overlapped background area from translational(top left) to affine(top right) and perspective(bottom)

2) Affine stage: To estimate of the camera zooming(s), rotation(φ_z) and finer tuning of translational components(g_x, g_y). This rigid affine model is the next finer approximation of the eight-parameter perspective transformation. The mapping between parameter (g_x, g_y, s, φ_z) and $(p_1, ..., p_6)$ can be obtain from following equation:

$$A(g_x, g_y, s, \varphi_z) = \begin{pmatrix} s\cos\varphi_z & -s\sin\varphi_z g_x \\ -s\sin\varphi_z & s\cos\varphi_z \; g_y \end{pmatrix} = \begin{pmatrix} p_1 p_2 p_3 \\ p_4 p_5 p_6 \end{pmatrix}. \quad (8)$$

3) Perspective stage: To estimate of (p_7, p_8) and fine tuning of $(p_1, ..., p_6)$. Difference from conventional MPEG-4 method, the gradient descent optimization is only repeated up to 5 times rather than 32.

1) Motion Ambiguity Caused by Rapid Camera Zooming: Theoretically, it is difficult to tell the difference between a translational motion and a zooming[22]. This motion ambiguity could cause GME inaccuracy, especially in the case of pure zooming. In order to solve this issue, zooming scale will be evaluated for zero camera moving. The corresponding GMC error between zooming and translational motion will be compared to determine the true motion.

2) Robust Aspect: The Object motion that is very different from camera movement usually has a negative influence on GME accuracy. In terms of GME, these objects may be considered as outliers[7]. In order to suppress the influences of outliers, in our proposed algorithm, a binary background mask indicating outlier location is generated at each stage; The outlier regions will not be included in the gradient descent energy minimization. Each global motion compensated(GMC) error frame is subdivided into blocks of size 4×4, and for each block, the sum of absolute differences (SAD) is computed. This size of the block is chosen because smaller blocks are

usually over-sensitive to image noise, while larger blocks may cause less detail. A block,B(*m,n*), with *m* being the horizontal and *n* being the vertical index, is used for gradient-based energy minimization only if

$$\sum_{m=1}^{4}\sum_{n=1}^{4}|I(4m+x_B,4n+y_B)-I'(x'(4m+x_B,4n+y_B),$$
$$y'(4m+x_B,4n+y_B))|<Th. \quad (9)$$

where (x_B, y_B) are the pixel coordinates within the block. Threshold *Th* is calculated adaptively in the following way: A histogram of global motion compensated error frame is constructed, then the value of *Th* is chosen such that 10 percent of the error frame pixels have a value larger than *Th*.

III. Testing Result and Discussion

Performance evaluation of our proposed method were compared to the conventional MPEG-4 using standard video sequences. The simulation sequences are: *Stefan*, *Soccer*,*Table Tennis*, *Bus* in CIF(352×288) and *Tennis* in SIF(352× 240). Both *Stefan* and *Soccer* show combined fast camera panning and zooming; and tennis has a large scale camera zoom-in to a table-tennis player.

Our algorithm source codes were written in MATLAB. We used mean absolute error(MAE) as a benchmark for our performance measurement. The MAE calculation is based only on the overlapped area in the GMC error frame. Both standard 30fps and 15fps sequences were tested for algorithm robustness and results are compared against that of conventional MPEG-4 method. For 30fps sequence the motion vector search range is [-32 +32] and [-16 +16] for x- and y-direction respectively, and the zooming scale search range is [0.97 1.03] with step increment of 0.01.

The testing results of sequences are shown in figure 8. Both 15fps and 30fps sequences were tested in forward/backward prediction to cover both camera zoom-in and zoom-out. Contract to the forward prediction, the anchor and target frame was exchanged with each other in backward prediction. Testing result has shown that both methods achieved the same accuracy when the zooming is within range [0.98 1.015]; beyond that range our proposed method is more convincing, for example, as shown figure 8d, the average MAE of our method in Soccer zooming section(no.226 to 264) is 5.8380 comparing against 10.7183 of conventional method; the zoom scale is [1.02 1.033] accordingly. In figure 8(e) the MAE difference can be as large as 5 because of the failure of conventional GME in camera zooming. Visual comparison of part result is also shown in GMC error frame of figure 9.

The minimum image registration does not always guarantee the same thing to the frame background. To ensure the optimal of proposed GME algorithm, frame background errors were also evaluated using manually segmented *Stefan*, *Table tennis* etc. In figure 8g, 8h and table, our method outperforms the conventional method in both background and entire frame error.

The algorithm efficiency of our method is also compared against the conventional MPEG-4 GME in figure 4. In figure, the starting MAE value our proposed method is picked right before affine gradient-descent stage, and that of reference method is before iterative stage. From figure 4, our proposed shows 8× faster than the reference method for that particular case. The rapid convergence of our method demonstrates that the affine stage is a good medium level approximation of camera motion, which leads to a better parameter optimization. Listed in table 1 and 2 are the sequence statistical results of both methods, again our method achieved better results compared with reference method.

In figure the average MAE values and the standard deviation of the entire sequences are listed in Table

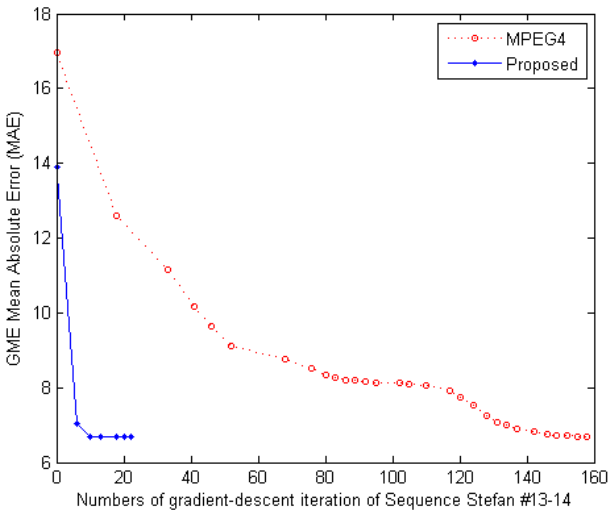


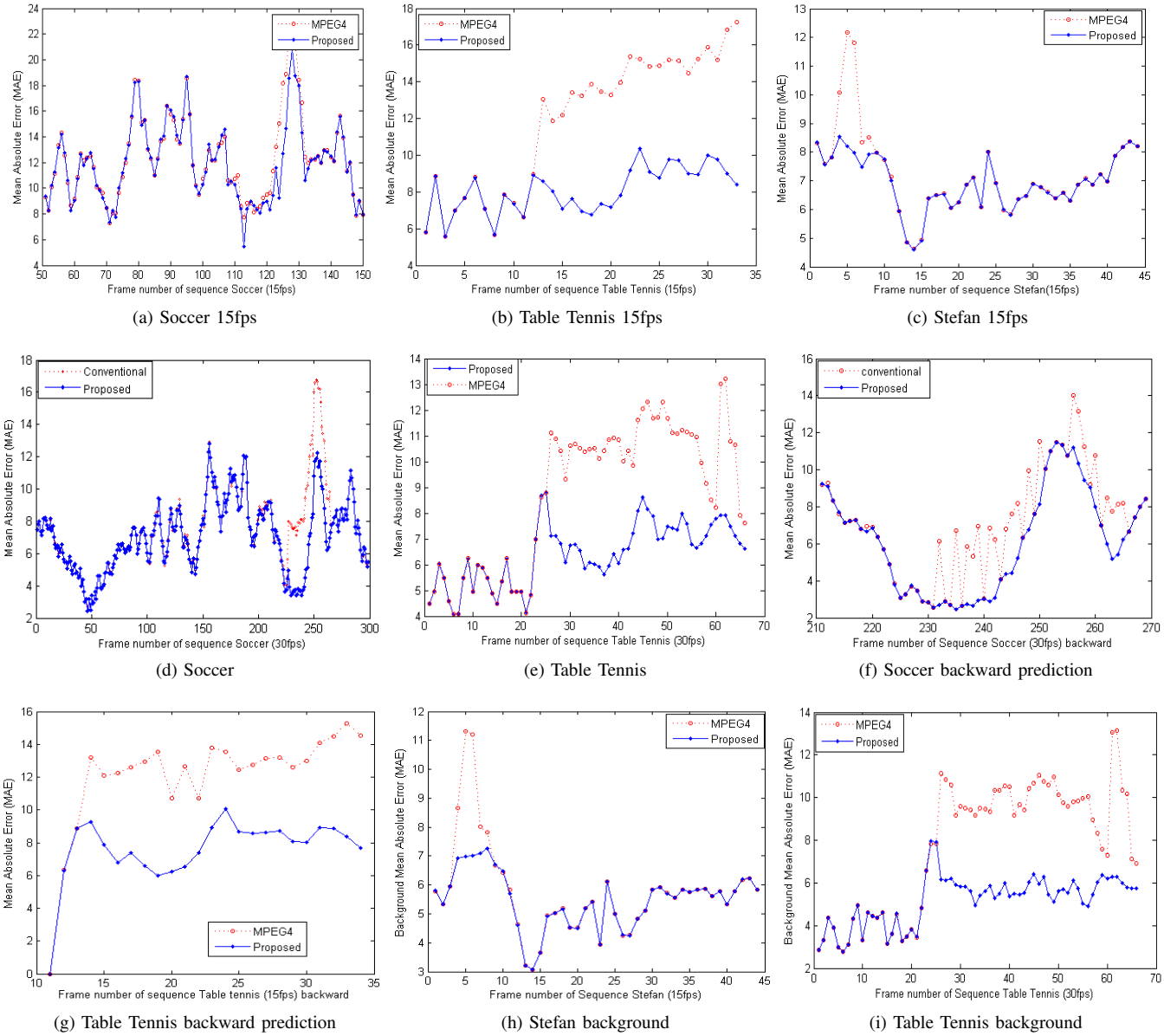
Fig. 5: Algorithm Efficiency Comparison

TABLE I: Average of mean absolute value of entire sequence

Sequence	Conventional method		Proposed method	
	MAE	STD	MAE	STD
Stefan	5.8272	0.989	5.8121	0.9717
Table Tennis	8.7015	2.8121	6.4182	1.1936
Soccer	7.6254	2.5459	7.0638	2.194
Bus(15fps)	10.6405	1.2432	10.4504	1.1403
Stefan(15fps)	7.1922	1.4955	7.0079	1.043
Soccer(15fps)	11.0942	3.5118	10.8417	3.2713
Table Tennis(15fps)	11.8493	3.7538	8.0791	1.3554

TABLE II: Average of background mean absolute value of entire sequence

Sequence	Conventional method		Proposed method	
	MAE	STD	MAE	STD
Foreman	3.5379	1.1953	3.4781	1.1059
Stefan	4.8367	1.0987	4.8222	1.0727
Table Tennis	7.7355	3.0347	6.4297	1.1972
Foreman(15fps)	4.5292	1.5468	4.4461	1.5026
Stefan(15fps)	5.7506	1.6359	5.4784	0.9849
Table Tennis(15fps)	11.8613	3.7546	8.0773	1.2804



(a) Soccer 15fps (b) Table Tennis 15fps (c) Stefan 15fps

(d) Soccer (e) Table Tennis (f) Soccer backward prediction

(g) Table Tennis backward prediction (h) Stefan background (i) Table Tennis background

Fig. 4: Sequence MAE(a-i), Sequence background MAE(h,i)

A. *Critical Factors of GME accuracy Analysis*

The accuracy and complexity of GME are strongly dependent on a series of specific properties of objects called critical factors[20].

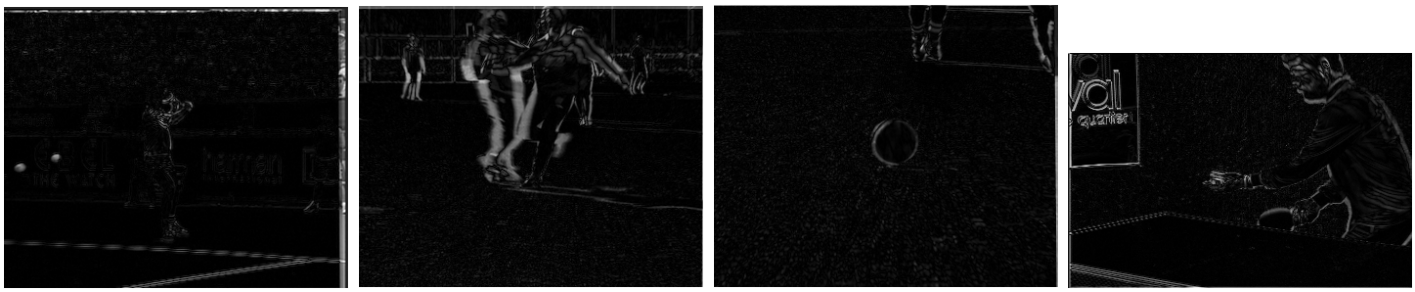
- Background Textural Complexity
  The reliability of GME is strongly relied on the textural complexity of background. The plain texture often make it hard to align the true background. As shown in figure 5, the alignment of the football players could lead to a lower error (MAE=12.88 versus 13.22) compared with that of true background matching; this kind of misalignment is the result of the background blurring when camera was undergoing a blunt moving.
- Dynamic background
  As shown in sequence *mobile*, the slow motion moving of the calender make it hard for GME to fully compensate the static background.
- Dominant Object Size
  As shown in sequence *Crew*, the moving object occupied more than 70 percent of total frame area, therefore the GME aligned the rightmost person rather than the slim background area.

B. *Future work*

Our proposed algorithm has achieved higher accuracy and efficient for consecutive video frames, however, it has difficulty in complicated case as described in the critical factors. Considering the continuity of camera motion, the temporal information could be used to improve the GME accuracy of entire sequence. The suggested improvement are described in following aspects:

1) To use a Kalman filter to verify the parameter reliability.

(a) Stefan (b) Soccer (c) Soccer (d) Table Tennis



(e) Stefan (f) Soccer (g) Soccer (h) Table Tennis

Fig. 6: Comparison of GMC error image: Proposed method(a,b,c,d); conventional MPEG-4 method(e,f,g,h), the lower gray level, the larger the overlapped area

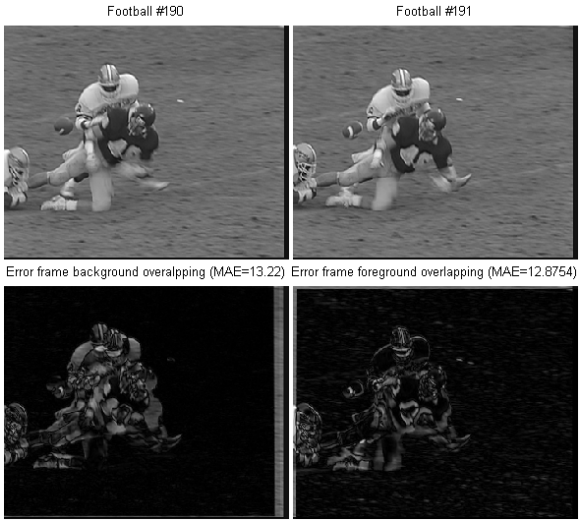


Fig. 7: Example of GME critical factor: Textural complexity (*Football*)

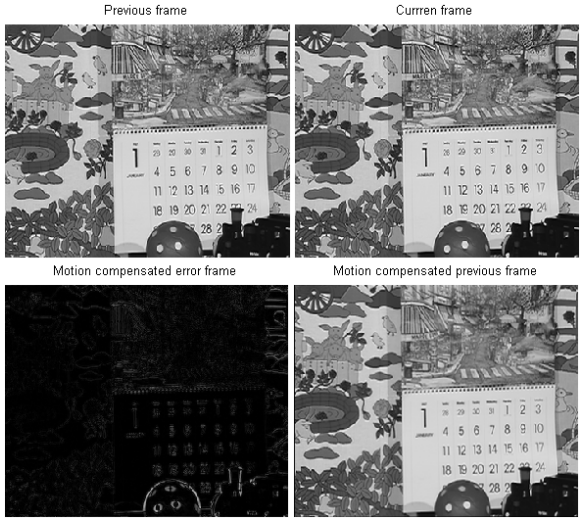


Fig. 8: Example of GME critical factor: Dynamic background (*Mobile*)

2) To use a background sprite technique to construct a frame background therefore to suppress the negative influence of moving objet in GME.
3) To use a motion predictor based on previous GMP to provide be a good initial guess for the next parameter[15][16].

IV. Conclusion

This paper proposed a fast and robust GME method in video sequences. This camera motion modeling is based on a coarse-to-fine approach with several improvements over the conventional MPEG-4 method.

First, An affine modeling stage is introduced between translational component estimation and perspective optimization. The introduced affine optimization is a medium level camera motion model, which brought down registration error faster and therefore provide a better starting point for following perspective optimization.

Second, a coarse camera zooming scale evaluation is introduced for the large camera zooming. This technique can be used to prevent false motion compensation caused by large scale zooming.

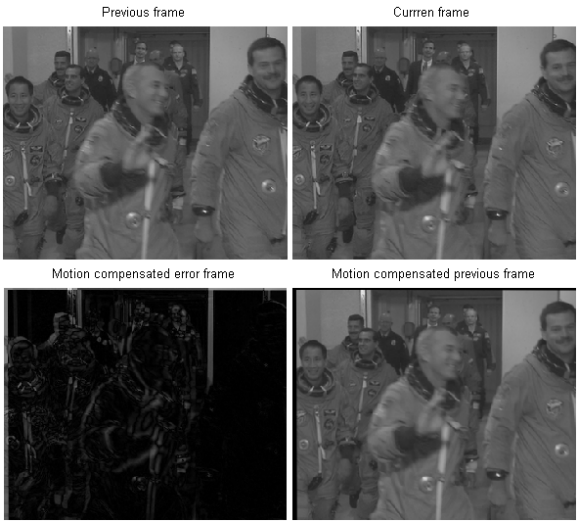


Fig. 9: Example of GME critical factor: dominant object size (*Crew*)

Third, our method has much less computational cost due to less number of gradient descent iterations compared with conventional MPEG-4 GME.

Overall, our proposed GME method has achieved a higher algorithm accuracy and efficiency.

REFERENCES

[1] Overview of the MPEG-4 Standard, V.18-singapre Version, ISO/IEC JTC11/SC29/WG11N4030, Mar. 2001.

[2] S.F. Chang, T. Sikora and A. Puri, Overview of the MPEG-7 standard, IEEE trans. Circuits Syst. Video Technol., vol. 11, no.6, pp.688-695, Jun. 2001.

[3] F. Dufaux and J. Konrad, Efficient, robust and fast global motion estimation for video coding, IEEE Trans. Image Process., vol. 9, no. 3,pp. 497501, Mar. 2000.

[4] D. Zhang and G. Lu., Segmentation of moving objects in image sequence: A Review, Circuits Systems and Signal Processing, Vol. 20 No.3, pp. 143-183, 2001.

[5] D. Capel and A. Zisserman. Automated mosaicing with super-resolution zoom, IEEE Int. Conf. on Computer Vision and Pattern Recognition, pp. 885- 891, June 1998.

[6] P. D. Capel. Image Mosaicing and Super-resolution. Springer, 2004.

[7] P. J. Huber. Robust estimation of a location parameter. Annals of Mathematical Statistics, (35):73 101, 1964.

[8] A. Krutz, M. Kunter, M. Mandal, M. Frater, and T. Sikora, Improved video coding using long-term global motion compensation," Visual Communications and Image Processing, vol. 5308, pp. 343- 354, Jan. 2004.

[9] Y. Chien, S.-Y. Ma, and Chen L.-G, Fast global motion estimation for MPEG-4 video compression In PACKET VIDEO, April 2003.

[10] R. L. Felip and L. Barcel, Robust Dominant Motion Estimation Using MPEG Information in Sport Sequences, IEEE Trans. Circuits Syst. Video Technol., Vol. 18, NO. 1. Jan. 2008

[11] E. Izquierdo, J. H. Xia, and R. Mech, A Generic video analysis and segmentation system, in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, May 2002, vol. 4, pp. 35923595.

[12] H. Xu, A. A. Younis, and M. R. Kabuka, Automatic moving object extraction for content-based applications, IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 6, pp. 796812, Jun. 2004.

[13] F. Moscheni, F. Dufaux, and M. Kunt, A new two-stage global/local motion estimation based on a background/foreground segmentation, in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, May 1995, vol. 4, pp. 22612264.

[14] C. Hsu and Y. Tsan, Mosaics of video sequences with moving objects, in Proc. IEEE Int. Conf. Image Processing, Oct. 2001, vol. 2, pp. 387390.

[15] W. C. Chan, O. C. Au, and M. F. Fu, Improved global motion estimation using prediction and early termination, in Proc. IEEE Int. Conf. Image Processing, Sep. 2002, vol. 2, pp. 285288.

[16] Y. He, B. Feng, S. Yang, and Y. Zhong, Fast global motion estimation for global motion compensation coding, in Proc. IEEE Int. Symp. Circuits and Systems, May 2001, vol. 2, pp. 233236.

[17] D. S. Farin, Automatic Video Segmentation Employing Object/Camera Modeling Techniques, PhD Thesis Holland 2005

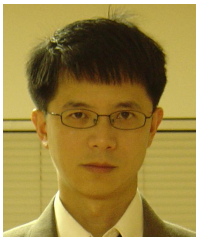
[18] X. Li, Multiple Global and Motion Models Used in Video Coding, PhD Thesis Georgia Institute of Technology May 2007

[19] H. Alzoubi and W. David Pan, Fast and accurate global motion estimation algorithm using pixel subsampling, Information Sciences 178 (2008)3415-3425

[20] F. Tiburzi and M. Escudero, A Ground Truth for Motion-based Video-object Segmentation, IEEE, 2008

[21] B. Qi and M. Ghazal, Robust Global Motion Estimation Oriented to Video Object Segmentation, IEEE Transactions On Image Processing,Vol.17,No.6,JUNE 2008

[22] L. Superiori and M. Rupp, Detection of Pan and Zoom in Soccer Sequence based on H.264/AVC Motion Information, WIAMIS 2009



Chun Qian has received the B.Sc. degree in applied physics from the Shanghai University, Shanghai, China, in 1989, and the M.Sc., degree in electrical and computer engineering from Michigan State University, East Lansing, MI, USA, in 2001. From 2003 to present, he has been in Doctorate program in Engineering. His research interests include global motion estimation, video object segmentation and fuzzy logic based artificial intelligence.

Performance Analysis of Routing Protocols for Wireless Ad-Hoc Networks

Sukhchandan Lally
Simon Fraser University
Vancouver, British Columbia, Canada
E-mail: lally@cs.sfu.ca

***Abstract*—Wireless ad-hoc networks are decentralized wireless networks that do not rely on an infrastructure, such as base stations or access points. Routing protocols in ad-hoc networks specify communication between routers and enable them to select routes between a source and a destination. The choice of the routes is performed by routing algorithms. In this paper, I use OPNET Modeler to simulate three routing protocols for wireless ad-hoc networks in several TCP and UDP scenarios. I analyze route discovery time, end-to-end delay download response time, and routing traffic overhead in static, less dynamic, and highly dynamic mobility scenarios. Simulation results indicate that the Ad-Hoc On-Demand Distance Vector (AODV) protocol is the most flexible when compared to Dynamic Source Routing (DSR) and Optimized Link State Routing (OLSR) protocols in the case of movement. Only the OLSR protocol meets the end-to-end delay requirements of less than 20 ms.**

***Index Terms*—AODV, DSR, MANET, OLSR, Wireless ad-hoc network**

I. INTRODUCTION

Wireless ad-hoc networks are a collection of mobile nodes that make up a multihop autonomous system. Their decentralized nature makes them suitable for various applications that cannot rely on a central node. These networks consist of multiple nodes and links. Each node requires a route for communication. Hence, each node participates in the routing process by forwarding data to other nodes.

In this paper, I describe a comparative performance analysis of three routing protocols (Ad-Hoc On-Demand Distance Vector (AODV), Dynamic Source Routing (DSR), and Optimized Link State Routing (OLSR)) for mobile ad-hoc networks (MANETs). In Section 2, I survey the literature. Section 3 describes simulated network topology while simulation scenarios are elaborated in Section 4. Section 5 provides results and finally, I conclude with Section 6.

II. AD-HOC ROUTING PROTOCOLS

Ad-hoc routing protocols control routing packets between computing devices in a mobile ad-hoc network. Nodes in ad-hoc networks are not aware of the network topology and have to discover it [14].

MANET routing protocols can be classified as unicast, multicast, and broadcast. The main goal of unicast protocols is to establish and maintain a route between a pair of nodes. They can be further subdivided into reactive (on-demand) and proactive (table-driven) routing protocols based on the method of acquiring information from the other nodes. In addition to these two main groups, there are hybrid routing protocols that combine the merits of both reactive and proactive routing protocols [1].

The advantage of on-demand routing protocols is that they generate less routing overhead compared to table-driven routing protocols. However, a source node may suffer from long delays to obtain a route to a specific destination [1]. The advantage of table-driven routing protocols is that a source node can obtain a routing path immediately if needed. However, they generate a high routing overhead [1]. A classification of ad-hoc routing protocols is given in Fig 1.

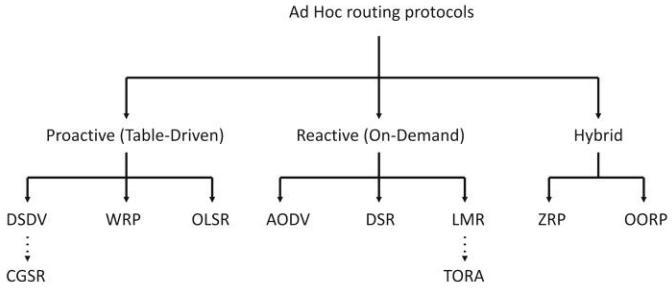


Figure 1.Ad-hoc routing protocols.

A. Ad-Hoc On-Demand Distance Vector (AODV)

Ad-Hoc On-Demand Distance Vector is one of the most popular reactive routing protocols. AODV routing algorithm is suitable for a dynamic self-starting network and users of ad-hoc networks. It ensures loop-free routes even while repairing broken links. Since the protocol does not require global periodic routing advertisements, the overall bandwidth needed for the mobile nodes is considerably smaller than in protocols that need such advertisements [1], [6].

AODV defines Route Requests (RREQs), Route Replies (RREPs), and Route Errors (RERRs) message types [1], [11]. These message types are received via User

Datagram Protocol (UDP). A source node initiates the Path Discovery operation and it broadcasts RREQ packet to its neighbors when it wishes to send a packet to a specific destination. This request is sent to all the neighboring nodes until either a neighboring node or destination node is located. During this process all the addresses from source to destination are recorded in the routing tables of the nodes. After RREQ reaches the destination or neighboring node, it replies with a RREP packet, and has a reverse path to the source. If there is any change in the network topology or there is a link breakage the node sends a RERR packet. Table 1 defines AODV parameters used in this project.

TABLE I. AODV parameters.

Attribute	Value
Route Discovery Parameters	Default
Active Route Timeout (seconds)	3
Hello Interval (seconds)	uniform (2, 2.1) uniform (10, 10.1)
Allowed Hello Loss	2
Net Diameter	16

B. Dynamic Source Routing (DSR)

Dynamic Source Routing is an on-demand routing protocol based on the concept of source routing where each routed packet carries in its header a complete and ordered list of nodes through which packet will traverse. Hence, intermediate nodes need not maintain up-to-date routing information in order to route the packets they forward [3], [12]. The protocol consists of two major phases: route discovery and route maintenance [3], [12].

When a source node wishes to send a packet to destination, it obtains a source route by the route discovery mechanism.

At first, a source node consults its route cache to determine whether it already has a route to the destination. If such a route leis not available, it initiates route discovery by broadcasting a Route Request packet. The Route Request packet is then answered by a Route Reply packet when the Route Request reaches either the destination or an intermediate node with an un-expired route to the destination in its route cache [3], [5], [12].

The route maintenance mechanism uses Route Error packets and acknowledgments. Route Error packets are generated to notify the source node that a source route is broken [3], [12].DSR parameters are defined in Table 2.

TABLE II. DSR parameters.

Attribute	Value
⊟ Route Cache Parameters	(...)
·· Max Cached Routes	Infinity
·· Route Expiry Timer (seconds)	60 300
⊞ Route Cache Export	Do Not Export
⊞ Send Buffer Parameters	Default
⊟ Route Discovery Parameters	(...)
·· Request Table Size (nodes)	16
·· Maximum Request Table Identifi...	16
·· Maximum Request Retransmissio...	16

C. Optimized Link State Routing (OSLR)

Optimized Link State Routing is a proactive routing protocol. It periodically exchanges topology information with other network nodes. The periodic nature of the protocol creates a large amount of overhead. OLSR reduces this overhead by using Multi Point Relays (MPR). Each node selects as MPRs a set of neighboring nodes and only those MPRs are responsible for forwarding routing messages and network-wide traffic. In OLSR, only nodes that have been selected as MPRs by a neighboring nodes announce this information periodically. Hence, the selected node has the ability to reach the node that has selected it as an MPR [1] [13].

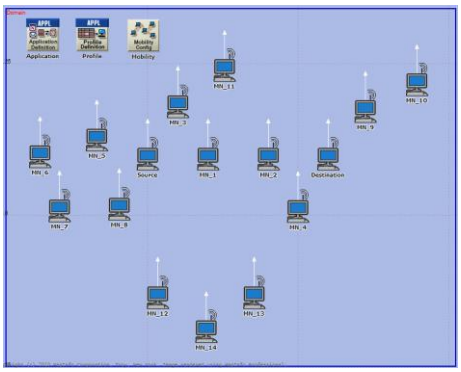
OLSR does not require reliable control message delivery because each node sends control messages periodically and, hence, can sustain reasonable loss of control messages. Each control message uses a sequence number, which is incremented for each message. Therefore, the protocol does not require sequenced delivery of messages [1], [13]. OLSR uses Topology Control (TC) messages to provide sufficient link state information to every network node to allow route calculation [1], [13].Table 3 defines OLSR parameters.

TABLE III. OLSR parameters.

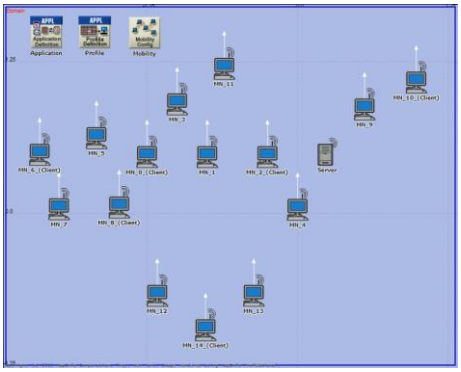
Attribute	Value
Willingness	Willingness Default
Hello Interval (seconds)	2 10 2 10
TC Interval (seconds)	5 5 25 25
Neighbor Hold Time (seconds)	6.0
Topology Hold Time (seconds)	15.0
Duplicate Message Hold Time (seconds)	30.0

III. OPNET SIMULATED NETWORK TOPOLOGIES

I have developed OPNET model of an ad-hoc network to evaluate the performance of three routing protocols in various environments.



(a)



(b)

Figure 2.Generic OPNET model: (a) UDP connection scenario and (b) TCP connection scenario.

Generic OPNET models for an ad-hoc network for highly dynamic environment with UDP and TCP connection scenarios are shown in Figure 2(a) and 2(b), respectively. Note that the routing protocol and mobility differ in each scenario. Every scenario consists of 16 wireless LAN (WLAN) nodes that are connected via low power wireless connection with transmission power of 5 mW and packet reception-power threshold of -90 dBm. Each node covers an area of approximately 675 m [4].Each node can only see its neighboring nodes because the distance between neighboring nodes is about 500 m.

A. Node Mobility

Mobility is one of the main attributes in ad-hoc networks. Modeling movement of a set of nodes is essential for evaluating the performance of an ad-hoc network. In this project, a random waypoint model available in OPNET Modeler was used.

I created three mobility models: static, less dynamic, and highly dynamic. In the static mobility model, I identified the best results for each routing protocol. In the less dynamic mobility model, I used very low-speed movement (walking speed) for a few mobile nodes. In the highly dynamic mobility model, I created high-speed movement (driving in the city) for all mobile nodes. I identified the routing protocol with the best performance in different dynamic environments and the routing protocol that is more stable in various mobility scenarios.

IV. SIMULATION SCENARIOS

I used the OPNET Modeler 16.0.A to simulate wireless ad-hoc networks with three routing protocols: AODV, DSR, and OLSR. I also analyzed the effect of periodic routing advertisement under different movement conditions.

The first scenario is a static scenario used to analyze a static ad-hoc network and to compare its performance with other scenarios. In the case of static network, nodes are motionless and, hence, all routes are valid for the entire simulation time.

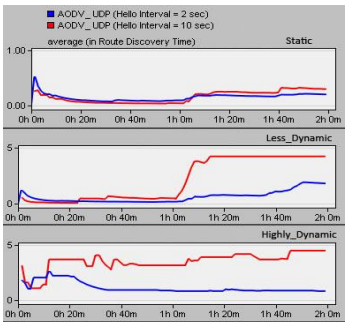
In the second scenario, some nodes moved with very low speed. The speed is comparable to the human walk (1 m/s). In this paper, this scenario is called as a less dynamic scenario.

The third scenario included high-speed nodes that moved with maximum speed equal to the speed of cars in a city (50 km/h). The nodes may or may not move during this simulation run.

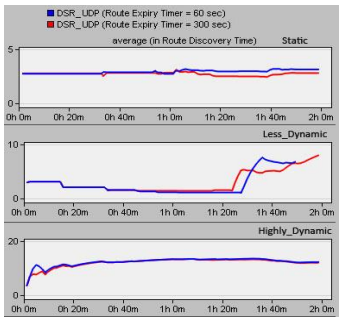
For each scenario, I considered two types of connections (UDP and TCP) and three ad-hoc routing protocols with various attributes. I created 48 simulation scenarios and analyzed route discovery time, end-to-end delay, download response time, and routing traffic overhead in TCP and UDP connection scenarios.

A. UDP connection OPNET model

In UDP connection scenarios, a two-hour interval of the Matrix III movie trace [10] was streamed from the source node to the destination node using various routing protocols and attributes. I created 24 simulation scenarios for UDP connection case. I chose only nine scenarios for the comparison analysis.



(a)



(b)

Figure 3.Average route discovery time in the UDP connection scenarios for (a) AODV (b) DSR.

The faster the nodes can find a route, the faster they can begin sending the video resulting in smaller end-to-end delay. Hence, to identify the best video streaming performance for a variety of AODV and DSR attributes, I examined route discovery time. Each scenario has different route discovery time, as shown in Figure 3(a) and 3(b). In case of AODV scenarios shown in Figure 3(a), AODV routing protocol with hello message interval of 2 seconds has better route discovery time as compared to other scenarios. In every DSR scenario shown in Figure 3(b), DSR routing protocol with route expiry timer of 300 seconds has better route discovery time.

OLSR is a proactive routing protocol and, hence, it has a route to the destination before it begins sending data. This protocol results in smaller end-to-end delay in streaming video packets in comparison to AODV and DSR.

Comparison of packet end-to-end delay in different scenarios with OLSR routing protocol is shown in Figure 4. The OLSR routing protocol with hello message interval of 2 s and topology control message interval of 5 s performs better in finding a route to the destination and in dealing with the movement.

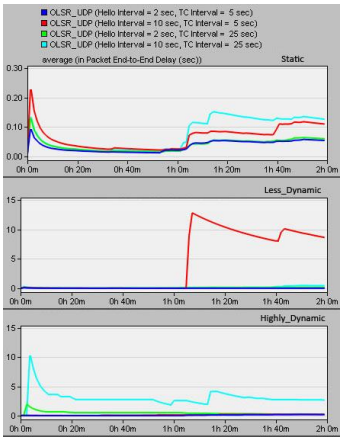


Figure 4.Average packet end-to-end delay in the UDP connection scenarios for OLSR.

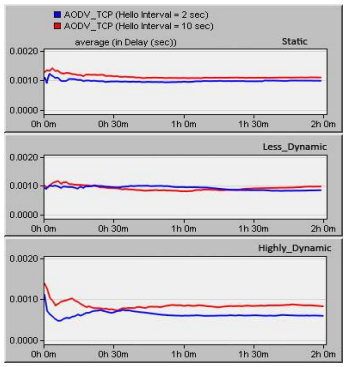
B. TCP connection OPNET model

TCP connection scenarios consist of 6 client nodes that download 50 KBytes of data from the server with different start times having uniform distribution of 20 s to 80 s. After 180 s, the client nodes restart downloading the data.

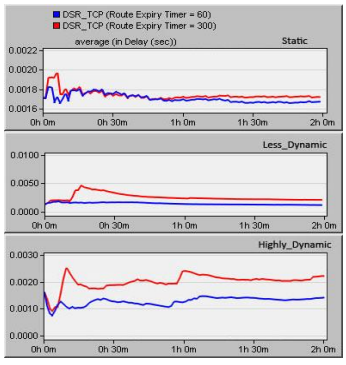
In the scenarios with the DSR routing protocol, I used two route expiry timers: 60 s and 300 s. In the first case, the route to the FTP server expires after 60 s and source node has to find a new route to send another file. In the second case, the route to the server remains valid and the client node may use the route to request another file.

I choose one scenario in each routing protocol scenario for comparison. Here, unlike the UDP connection scenario, I considered wireless delay.

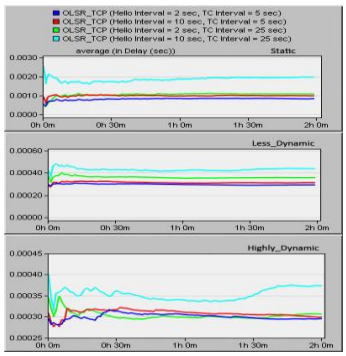
Average wireless delays for AODV, DSR, and OLSR are shown in Figure 5. AODV with hello message interval of 2 s has minimum wireless delay, as shown in Figure 5(a). The DSR with route expiry timer of 60 s has minimum wireless delay, as shown in Fugue 5(b). The OLSR routing protocol with hello message interval of 2 s and topology control message interval of 5 s has minimum wireless delay, as shown in Figure 5(c).



(a)



(b)



(c)

Figure 5.Average wireless delay in the TCP connection ad-hoc network for (a) AODV (b) DSR (c) OLSR.

V. RESULTS

In this Section, I compare routing protocols. I first consider route discovery time as one of the most important factors in on-demand routing protocols. I then consider packet end-to-end delay in UDP connection scenario streaming a video over the network, and download response time in TCP connection scenario downloading a file from an FTP server through several network nodes. Finally, I show the routing traffic overhead generated throughout the ad-hoc network employing each routing protocol.

A. Route Discovery Time

Route discovery time is an important factor in on-demand routing protocols. It causes large delays if the route discovery operation fails to find a route to the destination. Route discovery time for AODV and DSR routing protocols in UDP connection scenarios is shown in Figure 6. In the static UDP scenario, the route discovery phase in AODV is approximately 10 times faster than the route discovery phase of DSR in a static ad-hoc network. In case of less dynamic UDP network, AODV again has smaller route discovery time, which is similar to the highly dynamic scenario.

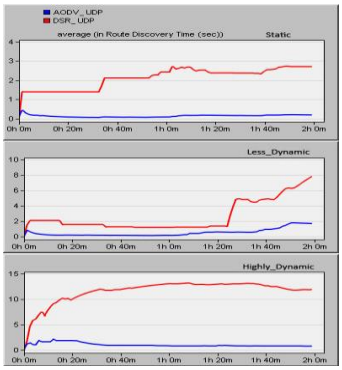


Figure 6.Average route discovery time (sec) in all UDP connection ad-hoc network scenarios for AODV and DSR.

The route discovery phase in AODV routing protocol is independent of the network topology, as shown in Figure 6. This implies that the changes in the network topology do not affect the performance of the route discovery phase of AODV in video streaming network. Unlike AODV, the route discovery phase in DSR depends on the network topology. The DSR route discovery time is higher in scenarios that include movements and the continuously changing topology, as shown in Figure 6.

Simulation results shown in Figure 7 indicate that for TCP connection scenarios, finding routes for AODV routing protocol are almost instantaneous while DSR routing protocol takes more time to discover routes.

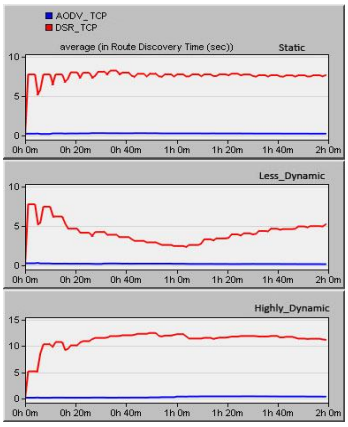


Figure 7.Average route discovery time (sec) in all TCP connection ad-hoc network scenarios for AODV and DSR.

B. End-to-end delay/Download response time

To find routes faster means to deliver data with less delay. The simulation results for UDP connection scenarios are shown in Figure 8. The end-to-end delay in the static network for all three routing protocols is approximately less than 0.5 s for most simulation scenarios. As I expected from the route discovery time, the AODV end-to-end delay is almost constant for all mobilities. However, the DSR end-to-end delay grows as mobility increases.

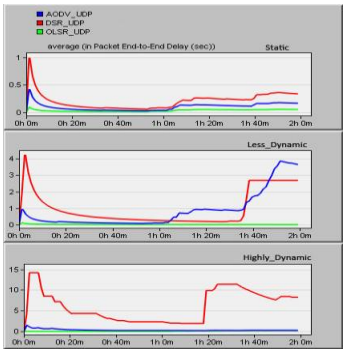


Figure 8.Average packet end-to-end delay (sec) in all UDP connection ad-hoc network scenarios for AODV, DSR, and OLSR.

The OLSR has the smallest delay in all mobility scenarios because it is a proactive routing protocol and it discovers routes before attempting to send any data. Hence, OLSR routing protocol has a general overview of network topology and every node in OLSR has at least one valid route to each reachable destination.

Download response time in TCP connection scenarios is shown in Figure 9. Although the route discovery time for DSR is higher than the AODV route discovery time, it has the smallest download response time. The DSR download response time is also smaller than download response time for OLSR.

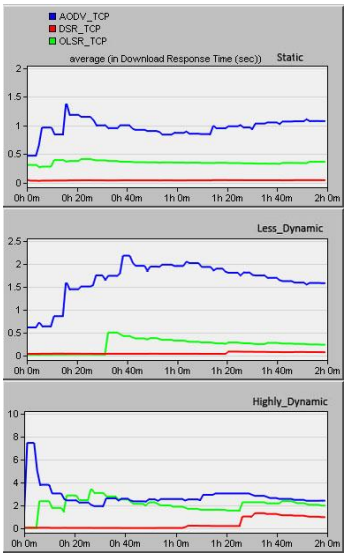


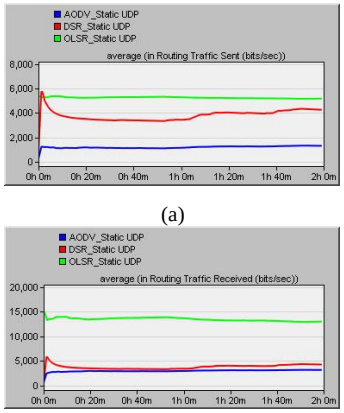
Figure 9.Average download response time (sec) in all TCP connection ad-hoc network scenarios for AODV, DSR, and OLSR.

The DSR has the smallest download response time, which might be due to the DSR source routing. The DSR uses source routing, which implies that a destination node does not need to discover a new route to the source node in order to send back the acknowledgement for each TCP packet.

C. Routing Traffic overhead

Routing traffic is another important factor in ad-hoc networks and may be a determinant factor in networks such as Wireless Sensor Networks (WSNs).

Average routing traffic sent and received throughout the UDP connection ad-hoc network for AODV, DSR, and OLSR is shown in Figure 10. As expected, OLSR has much larger routing traffic overhead in comparison to AODV and DSR. Figure 10 also illustrates an interesting result in OLSR routing. The protocol sends about 5,500 bits/second and receives about 14,000 bits/sec of routing traffic, which implies that each packet is duplicated approximately three times. Note that the sent and received routing traffic overhead is constant in DSR and is approximately the same in AODV.

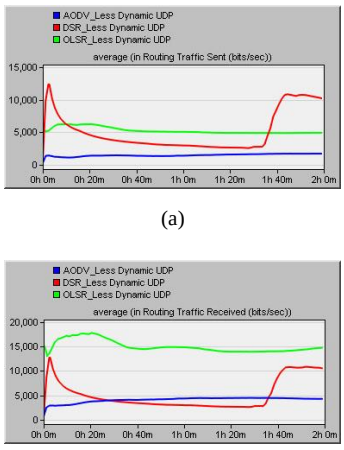


(a)

(b)

Figure 10.Average routing traffic sent (a) and received (b) in the static ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.

Results for less dynamic UDP connection ad-hoc network are shown in Figure 11. The only difference in routing traffic sent and received in less dynamic UDP connection network compared to static UDP connection network is a slight increase in routing traffic sent and received.

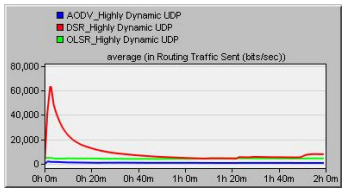


(a)

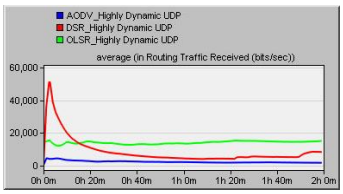
(b)

Figure 11.Average routing traffic sent (a) and received (b) in the less dynamic ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.

The DSR tendency to send more routing traffic throughout the network in comparison to static network is illustrated in Figure 11. The DSR sends more routing traffic in the presence of highly dynamic nodes and in video streaming scenario, as shown in Figure 12.



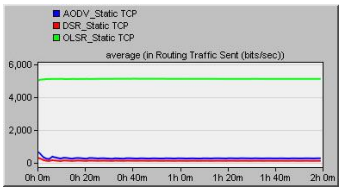
(a)



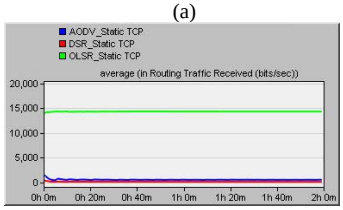
(b)

Figure 12.Average routing traffic sent (a) and received (b) in the highly dynamic ad-hoc network, UDP connection scenario for AODV, DSR, and OLSR.

Figures 13, Figure 14, and Figure 15 show the average routing traffic sent and received in the ad-hoc network in TCP connection scenario. These results show that although DSR routing traffic in video streaming scenario increases as nodes movement increase, DSR has consistent results in the file downloading scenario and it has the least amount of routing traffic compared to AODV and OLSR. Furthermore, OLSR shows the same results as UDP connection scenario and it has very large amount of traffic sent and received.

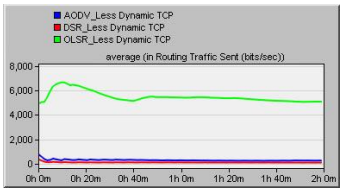


(a)

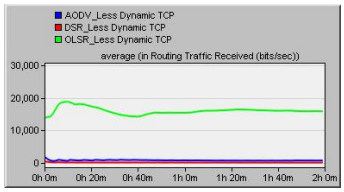


(b)

Figure 13.Average routing traffic sent (a) and received (b) in the static ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.

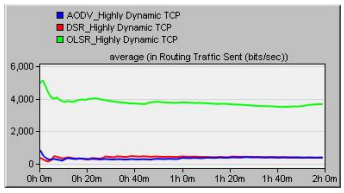


(a)

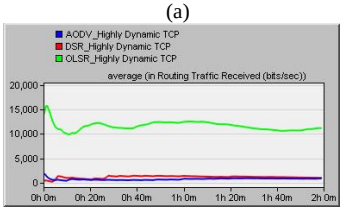


(b)

Figure 14.Average routing traffic sent (a) and received (b) in the less dynamic ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.



(a)



(b)

Figure 15.Average routing traffic sent (a) and received (b) in the highly dynamic ad-hoc network, TCP connection scenario for AODV, DSR, and OLSR.

VI. CONCLUSION

In this paper, I simulated and compared performance of various wireless ad-hoc routing protocols in various environments. In general, AODV is the most flexible protocol among the protocols that I considered and it works better in presence of movement while generating low routing traffic overhead.

AODV: Simulation results indicate that AODV is the most flexible routing protocol in the presence of movement. It also has the least routing traffic overhead in most scenarios and its routing discovery phase discovers routes very fast.

DSR: In video streaming scenarios, DSR does not exhibit a good performance, especially in presence of movement. In addition to large end-to-end delay in video streaming in DSR routing protocol, DSR also suffers from less flexibility in presence of movement and as the movement increases, the route discovery time and routing traffic overhead increase.

In case of TCP connection scenarios, DSR shows a good performance only in download response time and has low routing traffic overhead. However, DSR is inflexible in

video streaming scenarios. It can handle movement in the network used for transferring files.

OLSR: In the UDP connection simulation scenario with streaming a video over the network, only the OLSR routing protocol maintained the demand for end-to-end delay value less than 20 ms [10], while other protocols suffered from large end-to-end delay. The OLSR generates large routing traffic overhead to always have at least one route to every node in the network. In case of TCP connection scenarios, OLSR does not perform Ill. It generates a large amount of overhead to always know about network topology. It does not have smaller download response time than DSR.

In the presence of movement, DSR and OLSR impose large routing traffic overhead in ad-hoc network and may fail to perform in conditions with limited resources available (buffer, bandwidth, power.).

ACKNOWLEDGMENTS:

I wish to thank Dr. Ljiljana Trajkovic and Reza Qarehbaghi for their support in completing this paper.

REFERENCES

[1] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, *Ad Hoc Mobile Wireless Networks: Principles, Protocols, and Applications*, New York: Auerbach Publications, 2007.

[2] G. Jayakumar and G. Ganapathy, "Performance comparison of mobile ad-hoc network routing protocol," *IJCSNS International Journal of Computer Science and Network Security*, vol.7, no.11, pp. 77–84, Nov 2007.

[3] J. Broch, D. A. Maltz, D. B. Johnson, Y. Hu, and J. Jetcheva "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, Dallas, Texas, United States, Oct. 1998, pp. 85–97.

[4] M. T. Hyland, "Performance evaluation of ad hoc routing protocols in a swarm of autonomous unmanned aerial vehicles," M.Sc. Thesis, Air Force Institute of Technology, Ohio, USA, Mar. 2007.

[5] E. Royer and C. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Personal Communication Magazine*, vol. 6, pp. 46-55, Apr. 1999.

[6] A. Suresh, "Performance analysis of ad hoc on-demand distance vector routing (AODV) using OPNET simulator," M.Sc. Mini Project, University of Bremen, Bremen, Germany, 2005.

[7] K. Gorantala, "Routing protocols in mobile ad-hoc networks," M.Sc. Thesis, Umeå University, SIden, 2006.

[8] Cisco Systems, Inc., *Cisco Network Planning Solution Standard Models User Guide*, Cisco Systems, Inc., 2005.

[9] OPNET Technologies, Inc. Making Networks and Applications Perform, "HOW TO: Design Mobile Ad Hoc Networks and Protocols," OPNET Technologies, Inc., 2007.

[10] W. Hrudey and Lj. Trajkovic, "Streaming video content over IEEE 802.16/WiMAX broadband access," *OPNETWORK 2008*, Washington, DC, Aug. 2008.

[11] Ad hoc On-Demand Distance Vector (AODV) Routing [online]. Available: http://www.ietf.org/rfc/rfc3561.txt.

[12] The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4 [online]. Available: http://www.ietf.org/rfc/rfc4728.txt.

[13] Optimized Link State Routing Protocol (OLSR) [online]. Available: http://www.ietf.org/rfc/rfc3626.txt.

[14] List of ad-hoc routing protocols [online]. Available: http://en.wikipedia.org/wiki/List_of_ad-hoc_routing_protocols.

BIOGRAPHY:

Sukhchandan Lally received her bachelors of technology in electronics and communication engineering from Lovely Institute of Technology, Punjab Technical University, India in 2008.She is currently undergoing her masters of applied science from Simon Fraser University, Canada. Her research interests include communication networks.

Identifying the Most Useful Papers to Read from the Disciplinary Literature

Shabnam Shariaty
School of Computing Science
Simon Fraser University, BC, Canada
{shabnams}@cs.sfu.ca

Due to the huge number of papers published in any research area, developing familiarity with a new area is an exhausting and time consuming task for any researcher. On the other hand, a paper may present the general ideas of some of its referenced papers as well. In this paper, we first develop a method for analyzing this relationship between papers. Then, we propose an algorithm for leveraging these relationships and identifying a number of key papers out of a given literature which contain the most useful information about major advances in that research area. We evaluate our analysis and algorithm by comparing our results with the suggestions of human judges, which confirms the accuracy and usefulness of our method.

Categories and Subject Descriptors: H.1.0 [**Models and Principles**]: General

General Terms: Algorithms, Design

Additional Key Words and Phrases: Content coverage, key paper, literature review

1. INTRODUCTION

The most common problem for every researcher who intends to become familiar with a new research domain is confronting a huge number of papers that need to be examined. However, by reading each paper, one can often obtain useful information about the general ideas of some of its referenced papers. The extent of this property, to which we refer as the *content coverage relationship* between two papers, can significantly vary from paper to paper. By discovering and taking advantage of these relationships between papers, we would like to identify a set of key papers in the new research area, which cover the most useful information about major advances in the field.

Due to its importance, the problem of automatically summarizing a research domain has received significant attention, and several approaches have been previously proposed in the literature. However, most of this work focuses on generating per-paper text summaries [Qazvinian and Radev 2008], [Mohammad 2009], [Qazvinian and Radev 2010]. Our work, on the other hand, is aimed at identifying a number of key papers from a given literature such that they best cover and/or point to the general ideas proposed in the literature.

In particular, we first propose a method to analyze the content coverage relationship between a paper and its references; note that such coverage relationship between papers A and B does not necessarily indicate that reading paper A would eliminate the need to read B (i.e., the cases where paper A briefly points out the main idea of B would also count). According to these relationships, we model the

given literature with a directed graph, on which we formally define our problem. We then prove that this problem is intractable (NP-Complete). Thus, we solve the problem through an approximation algorithm which leverages the classic solution of a well-known problem in graph theory called the Dominating Set problem. Finally, we evaluate our method for analyzing the content coverage relationship between papers and our algorithm for identifying key papers by comparing our results in with the opinions of human annotators for both steps, which confirms the accuracy and usefulness of the proposed solutions.

The rest of this paper is organized as follows. In Section 2, we formally define our problem and present our solutions. The dataset we use for our experiments in this paper is briefly introduced in Section 3. In Section 4, we evaluate the proposed solution, and Section 5 concludes the paper.

2. THE PROPOSED ALGORITHM FOR IDENTIFYING KEY PAPERS

In this section, we present the proposed algorithm for solving our problem. The input to our problem is a corpus of papers, including their contents and references to each other. The output is a set of key papers that well represent the area. Our solution to this problem consists of two steps: modeling the corpus with a directed graph, and then solving the problem through the solution to a known problem in graph theory called the Dominating Set problem.

2.1 Modeling the Problem

In order to identify a number of papers that have the best coverage of what has been proposed in a research area (i.e., the corpus) we need to obtain a clear picture of how papers include each others' content, out of the texts of the papers and their citations to each other. Given a data set containing this information, we model the set of papers with a weighted directed graph, where each node represents a paper, and an edge (u, v) indicates that paper u cites paper v. The weight of each edge (u, v) represents how well the information of paper v is reviewed by paper u, meaning that by reading paper u, one can obtain useful information about what paper v proposes.

The first step to solve our problem is to construct this graph and determine the weight of its edges. The intuitive idea for weighting the edges is as follows. Consider a k-tuple of words (e.g., $k = 2$) repeated multiple times in paper v. In this case, the tuple is either a common combination in English, such as "given that", or possibly a key term specific to that paper. We suggest to distinguish these two cases by comparing the number of occurrences of this k-tuple in papers u and v: multiple occurrences in paper v and very few (and at least once) in paper u, while paper u has cited paper v, is a reasonable indication that the tuple is most likely a key term in paper v, and that paper u talks about paper v.

Accordingly, we define the weight of edge (u, v) as follows:

$$w(u, v) = \sum_{x \in T(u) \cap T(v)} \frac{1}{n_u(x)} (1 - \frac{1}{n_v(x)}), \tag{1}$$

where x is a k-tuple, $T(u)$ is the text of the paper represented by node u, and $n_u(x)$ is the number of occurrences of x in paper u. These weights indicate the accuracy

of the hypothesis that paper u reviews paper v.

We then convert our graph to a simple unweighted graph by removing the edges that have a weight less than a given threshold α (discussed shortly), and ignoring the weights of the remaining edges.

2.2 Obtaining the Solution Set

Having created the desired graph, we need to obtain a subset of papers with a size of, at most, s papers which has the most coverage of the corpus; s is assigned a value based on the total number of papers and the preferences of users. Obtaining the aforementioned subset of nodes, which is named a Dominating Set, is a well-known problem in graph theory. More specifically, a Dominating Set for a graph $G = (V, E)$ is a subset $D \subset V$ such that every vertex v not in D is joined to at least one vertex u in D by an edge. In case of a directed graph, this edge must be from u to v, not vice versa. Given a graph G and an integer s, the Dominating Set problem consists of identifying a subset of, at most, s nodes in G with the maximum coverage. This problem is known to be a classic NP-complete problem.

For the Dominating Set problem on a graph of size n nodes, there is an approximation algorithm with ratio $log(n)$, meaning that its result in the worst case is guaranteed to be at most $log(n)$ times the minimal solution. This approximation algorithm is based on the approximation solution to the Set Covering problem; the Dominating Set problem is an special case of the Set Covering problem. The Set Covering problem is defined as follows: Given a universe U of n elements and a collection of its subsets, which may overlap, obtain a minimum number of these subsets such that the union of them contains all elements in the universe.

The Dominating Set problem [Melnikov 1998] for undirected graphs can be converted to the Set Covering problem [Vazirani 2003] as follows. The universe U for the Set Covering problem instance consists of V (the set of nodes) in the graph of the Dominating Set problem, and each subset of the aforementioned collection represents a node along with its neighbors.

Since we have modeled our problem with a *directed* graph, we need to modify the above procedure so that we can use the approximation algorithm of the Set Covering problem. That is, by each subset in the Set Covering problem instance, we represent a node v along with only those neighbors of v that have an incoming edge from v.

The approximation algorithm of the Set Covering problem that we will apply for our directed Dominating Set problem is as follows. By a greedy approach, in each step we select the subset that has the most number of members in the universe, and then we remove those members from the universe. We continue these iterations until the universe becomes empty (i.e., it is fully covered by the selected subsets). At this point, the graph nodes represented by these subsets constitute the answer set to our original Dominating Set problem.

3. THE DATASET

We extract our dataset from the ACL Anthology Network (AAN) [1]. ANN is a collection of more than 15,000 papers over a period of four decades. All papers are

[1] `http://clair.si.umich.edu/clair/anthology/`

Table I. The number of content coverage relationships between pairs of cited papers.

Judges \ Us	Yes	No
Yes	63	15
No	19	135

from the ACL Anthology published in the Computational Linguistics journal and proceedings from ACL conferences and workshops, so they more or less belong to the same broad domain and are suitable for our purpose. AAN includes the texts of the papers and their citation network in the ACL Anthology.

Out of this broad set of papers, we select a subset of 2,500 by keeping only the papers that contain the word "summ" and filtering out the rest, which provides us with a set of papers concentrated around a more specific domain: summarization. Also from the list of citations of papers to each other, we only keep those whose both referenced and referencing papers are within the set of 2,500.

In order to extract k-tuples from the papers and count their frequencies, we use the CMU-Cambridge Statistical Language Modeling Toolkit[2].

4. EXPERIMENTAL RESULTS

In this section, we evaluate our proposed solution using the dataset described in section 3. We compare our model (i.e., how well the edges of our graph represent content coverage) as well as the final answer set obtained by our approximation algorithm against the answers suggested by human judges.

To collect the reports of judges, we select a random subset of 500 papers out of our total 2500 papers, and for each paper, we ask 2 human annotators to manually analyze its citations that are within the 500-paper subset. In our experiments, this approach has resulted in 150–250 citations, each of which is a potential candidate for the content coverage relationship to be analyzed by the judges. For detecting such relationships by our algorithm, we have used a value of $k = 2$ in our experiments, meaning that we analyze 2-tuples as potential key phrases for indicating content coverage.

Recall that in order to detect the coverage relationship between two papers, we use a threshold α to cut weak edges. Because this threshold needs to be set according to some experts' opinions, in here we use the opinions of annotators themselves to set α. Note that this method is only used for setting the input parameter α determining the number of coverage relationships to report; this method enables us to compare the N relationships reported by annotators against the top N relationships reported by our model. Thus, we experimentally set α such that the number of edges we consider would be equal to the number of coverage relationships suggested by annotators.

Table I illustrates the comparison between the result reported by our model and the ones by the annotators for each citation link between two papers. As the figure confirms, the overlap between our reports and the ones by the annotators is 85%. The false negative ratio (i.e., the proportion of actual content coverage relationship that we identified as non-covering citations) is 8%. The false positive ratio (i.e.,

[2]http://mi.eng.cam.ac.uk/~prc14/toolkit.html

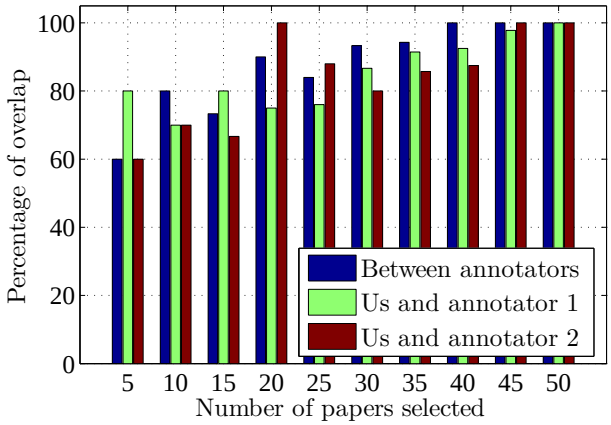


Fig. 1. Comparison between the top $5, 10, \ldots, 50$ papers suggested by our algorithm and by annotators.

the proportion of non-covering citations that we identified as content coverage relationship) is 6%. As these results confirm, our algorithm has an acceptably high accuracy.

Next, we evaluate the proposed approximation algorithm. Ideally, we would like to ask the annotators to identify a ranked subset of M papers (e.g., $M = 50$) out of the total 2500 papers that include the most from the research domain, but it is infeasible. Thus, we run our algorithm for $s = M$ to obtain a subset of M papers that we expect to be the best covering ones, and we ask the annotators to read only these M papers and rank them according to their coverage of other papers. This ranking is done with a granularity of 5 papers, meaning that the annotators report 5 papers as the best covering 5 (no ranking within these 5), another 5 papers as the second best, and so on. We then compare this ranked list with the one calculated by our algorithm, which we can be obtained by running our algorithm gradually for $s = 5, 10, \ldots, M$.

Figure 1 illustrates the result of this comparison, that is, how close our suggestions for the top x papers are to the top x papers suggested by each of the two annotators. We also depict the overlap between the suggestions of the two annotators to highlight that there is no exact answer to the problem of identifying the top x papers, due to the subjective nature of the problem. As Figure 1 depicts, on average there is about 80% of overlap between our results and those of annotators, which is more or less the same as the overlap between the opinions of the two annotators themselves. The percentage of overlap is less for smaller numbers of papers ($x \leq 15$), where a few mismatches result in a greater drop in the overlap percentage. In summary, Figure 1 confirms that our algorithm has a reasonably high accuracy because the set of papers suggested by our algorithm are close to the ones suggested by human judges.

5. CONCLUSION

In this paper, we have studied the problem of identifying the most useful research papers to read out of a large literature of works in a given research domain. These papers are selected in a way that they best include (and point to other papers containing) the main ideas proposed in the literature. For this purpose, we defined a content coverage relationship between two papers where one cites the other, which is detected in our model by analyzing how the key phrases of the referenced papers appear in the referencing papers. Based on these relationships, we present an algorithm for identifying the most useful s papers where s is an input parameter determining the desired number of papers to suggest. We evaluate the proposed model and algorithm by comparing our results against the suggestions of human judges, and we demonstrate that it has a high level of accuracy for identifying the most useful papers.

REFERENCES

Melnikov, O. 1998. *Exercises in graph theory.* Mathematics. Springer.

Mohammad, S. 2009. Using citations to generate surveys of scientific paradigms. In *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT'09)*. Boulder, CO, 584–592.

Qazvinian, V. and Radev, D. 2008. Scientific paper summarization using citation summary networks. In *Proceedings of the International Conference on Computational Linguistics (COLING'08)*. Vol. 1. Manchester, UK, 689–696.

Qazvinian, V. and Radev, D. 2010. Citation summarization through keyphrase extraction. In *Proceedings of the International Conference on Computational Linguistics (COLING'10)*. Beijing, China. Accepted to appear.

Vazirani, V. 2003. *Approximation Algorithms.* Springer-Verlag.