Measurement Analysis of IP-based Process Control Networks

Young J. Won[1], Mi-Jung Choi[1], Myung-Sup Kim[2], Hong-Sun Noh[3], Jun Hyub Lee[3], Hwa Won Hwang[4] and James Won-Ki Hong[1]

[1] Dept. of Computer Science and Engineering, POSTECH, Korea
{yjwon, mjchoi, jkbon, jwkhong}@postech.ac.kr

[2] Dept. of Computer and Information Science, Korea University, Korea
tmskim@korea.ac.kr

[3] Electric &Control Maintenance Dept., POSCO, Korea
[4]Technical Research Laboratories, POSCO, Korea
{vishnu4, mujigae, hwawon}@posco.co.kr

Abstract. This paper presents a measurement study of the traffic traces from the industrial process control IP networks. We present some interesting and unique traffic characteristics of the IP networks which support the control of manufacturing and precision-control machines. Understanding their traffic behaviors would help us to operate the fault-tolerant control IP networks, where the cost of network malfunctioning is far more severe than ordinary IP data networks. We observe rather steady and cyclic traffic patterns in the collected traces between the control IP network entities, mainly PLCs and process controllers.

Keywords: passive measurement, traffic analysis, process control IP networks

1 Introduction

Process control networks must sustain the robust communications between controlling devices and controlled devices in a manufacturing environment. There are several categories of control networks, namely Building Automation (BA), Factory Automation (FA), and Process Automation (PA) [1]. These networks are deployed in mission critical operations which require a maximum level of network stability. However, we know little about the traffic behavior of such networks, simply because we are dealing with very secure and closed networks which entail company's confidentiality.

Existing process control network technologies (e.g., FOUNDATION Fieldbus [3], PROFIBUS [4], MODBUS [5], BACnet [6], Lon Works [7], etc.) are developed separately from the relatively recent emergence of Ethernet and IP technology. The newer versions of them have adapted Ethernet (e.g., Industrial Ethernet) and IP for low cost, high scalability, and easy maintenance purposes. A few studies [1, 2] have introduced some important issues of moving towards the all IP-based control networks, such as QoS requirements of the control network and existing IP network security concerns. Nonetheless, the decision of such change in a manufacturing plant

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment)" (IITA-2006-C1090-0603-0045)

environment is beyond the technical superiority of Ethernet and IP because it involves a huge investment.

Unlike IP data networks (e.g., Internet, enterprise networks), the traffic characteristics of control IP networks have not been studied in-depth previously. Due to significant differences of traffic nature, the general OA&M guideline running IP data networks may not coincide fully with process control IP networks. Understanding the traffic behavior helps us to operate fault-tolerant control IP networks where the cost of network malfunctioning is far more severe than the IP data networks. We present a measurement study of the traffic traces from the real industrial process control IP networks. To our knowledge, this is the first work to provide an empirical traffic analysis in such networks.

The remainder of this paper is organized as follows. Section 2 provides an overview of our measurement environment. In Section 3, we present the results of empirical analysis of the collected traffic traces. Finally, we summarize our findings and discuss possible future work in Section 4.

2 Measurement Environment

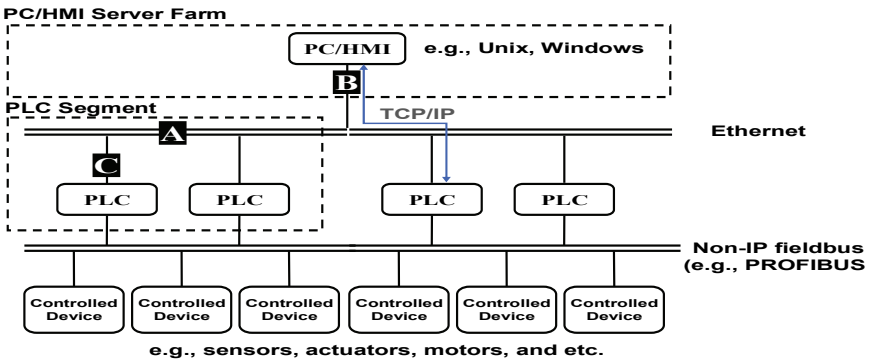

Figure 1. Control IP network topology & Measurement points – A, B, and C

A typical process control network is illustrated in Figure 1. We have captured traffic traces from three different measurement points (illustrated as A, B and C in Figure 1). The process control network elements follow a hierarchical model where the controller at the top triggers actions in one or more controlled devices. The following are brief descriptions of the component elements and their role.

- *Process Controller (PC) – This is a part of the software and hardware package provided by the PLC vendors. It is process control software on a computer running UNIX or Windows that can remotely access PLCs. Custom built or vendor provided server applications are placed in a PC to communicate with PLCs. It also communicates with the machines running Human Machine Interface (HMI) solutions which provide graphical presentation of real-time process status monitoring. The communication with PLCs is established over reliable TCP/IP.*
- *Programmable Logic Controller (PLC) - It is a microprocessor computer for process control attached to a process control network. A complex sequence*

control of machinery (or low end controlled devices on factory assembly line) is handled by the custom-built software programs running in PLC. The PCLs in our monitoring environment are equipped with two separate interfaces: Ethernet (e.g., RJ-45) and PROFIBUS (e.g., EIA-485).

- *Controlled Devices – These machineries refer to sensors, actuators, motors, etc. They receive the command signal from PLCs via the embedded interface.*

Two measurement domains exist in process control networks: PC to PLC network and PLC to controlled device network. It is a combination of IP-based and non-IP based control technologies; in fact, PLC acts as a gateway accessing the lower end devices, which are on non-IP based networks. In this study, we focus on the first half of the networks (i.e., PC to PLC). A PLC segment refers to a group of PLC networks at the edge.

We have collected the traffic traces from a various points of the process control networks using the standard libpcap [8]. Points A, B, and B in Figure 1 refers to the top of PLC segment, the process control backbone network, and the nearby end-host (PLC device), respectively. These representative locations are carefully selected to provide a precise and unbiased snapshot of the network.

3 Empirical Data Analysis

The experimental data set we used was collected at the process control networks of POSCO, the world's fourth largest iron and steel manufacturer [11]. It operates a number of plants world-wide and a single operational site consists of about 40 manufacturing plants, equally 40 process control networks where they are organized in a synchronous and sequential order. Each process control network is a group of edge network segments. In simpler terms, a number of networks are working together to interconnect the machineries running continuously on a conveyor belt.

Our assumption is that no other data traffic is injected into the monitoring PLC networks. This was achieved by the complete isolation of PLC network from the Internet or any other enterprise networks. The private IPs and dedicated Ethernet links were assigned to PLC devices.

3.1 Traffic Summary

Table I illustrates the traffic summary of the collected traces. We have analyzed a week-long trace at one of the edge segments, a typical working-hour trace at the plant backbone, and short (e.g., 5 minutes) traffic traces at the end-host segment – Segment A, B, and C, respectively. The actual name of each process segment is undisclosed due to security reasons. Most traffic is exchanged using TCP and the average utilizations yield a very low percentile in the 100 Mbps physical links environment. We observed the total traffic volume of Segment B regardless of the three times much as the packet counts of Segment B. It implies a low yielding traffic volume and a major occupancy of light-size packets. Interestingly, in Segment B, only 533 flows are TCP flows and responsible for 121M packets. In fact, a small number of identical sessions with fixed amount of hosts (PLC and PC) are continuously observed and generate the traffic.

Table I. Summary of the datasets

Set	Date	Duration	Byte	Packets	Flows	TCP%	Util.%	Location
Segment-A	2006-09-29	170 hr (7 days)	63.5 GB	542 M	48 K	98	1	Edge
Segment-B	2007-02-27	10 hr (12:00~20:00)	74 GB	122M	25 K	99	19	Backbone
Segment-C	2006-05-11	5 min (13:15~13:19)	22 MB	84 K	48	99	0.57	Edge

Figure 2 (a) illustrates a long-term traffic pattern of particular PLC segment – Segment A. The graph shows that the bandwidth consumption is not bound to the time-of-day effect in the typical IP data networks where more traffic is generated during the day or working-hours. Its bandwidth usage is very steady and predictable throughout the course of monitoring period. A few sudden drops in the graph, the shade regions, are observed which indicates an instant shutdown of the process due to scheduled or unscheduled maintenance purpose. Its bandwidth consumption is strictly proportional to the number of devices in the network. Thus, the network planning for control IP network can rely on a very precise projection of bandwidth growth model which is almost impossible in other type of IP networks.

The Segment B traces are a collection of multiple instances of the Segment A level of traffic. Figure 2 (b) shows a short-term bandwidth measurement, but much larger traffic volume at the backbone. Its pattern closely coincides with the behavior as in Figure 2 (a). Indeed, the microscopic view of traffic behavior can reflect the long-term behavior without loss of generality in control IP networks because there exists a fixed pattern in every session occupying the control IP networks. More details will be covered in the following sections.

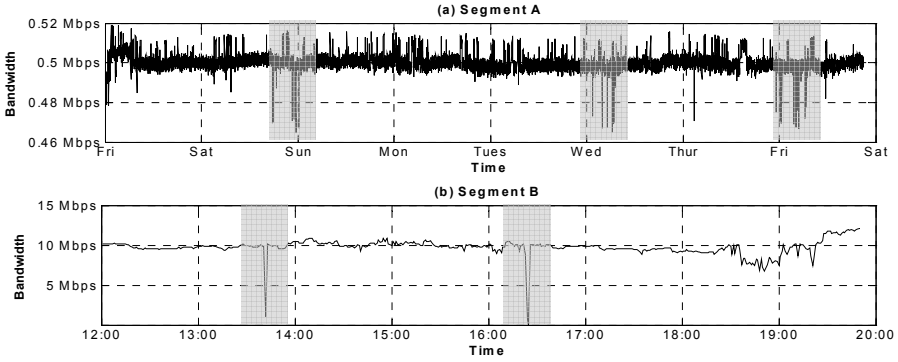

Figure 2. Traffic Volume – Segment A & B

3.2 Traffic Cycle

Figure 3 illustrates the four representative packet arrival patterns in PC-PLC sessions. The upper plane (above 0 on y-axis) of the graph indicates a unidirectional packet size arrival sequence according to the packet inter-arrival time. The lower plane is a packet size arrival sequence of the corresponding reverse transmission. Thus, a single graph represents two bidirectional flows. The packet inter-arrival time is measured in time granularity of millisecond. The dense region of graph implies that

the packet inter-arrival gap is reduced. In all four graphs, we observe a unique and regular cycle of dense and sparse region occurrences over a short time period as well as the packet size distributions. For example, Figure 3 (a) has about 12-second cycle of two dense regions followed by a sparse region in both upper and lower planes. In the lower plane, it also shows a regular cycle of 1000 bytes packet transmission in every 1.5 second. In a similar fashion, the rest of the graphs can be expressed as a traffic pattern candidate for general PLC-PC sessions. All the sessions in our measurement belong to one of the pattern shapes shown in the graphs of Figure 3. Figure 3 (b), (c), and (d) illustrated the session traffic patterns for different PLC vendor solutions or sessions involving possibly in different processes.

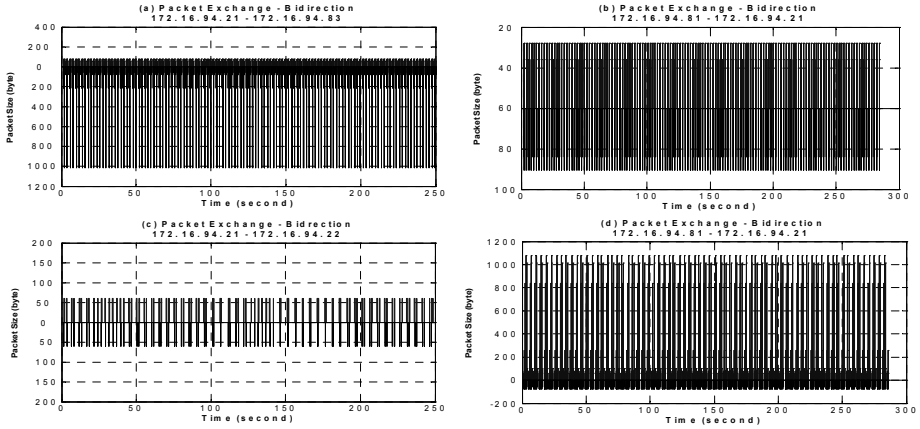

Figure 3. Bidirectional packet-size arrival sequence patterns of PC-PLC sessions

The average inter-arrival time of these four sessions ranges from 120 ms to 1.5 s. Note that, the selected sessions were operational without any performance irregularities at the time of monitoring period. A matter of hundreds of millisecond or above may not be acceptable a set of values for a packet delay in the IP data networks. It is a unique characteristic where relatively longer packet delays are acceptable.

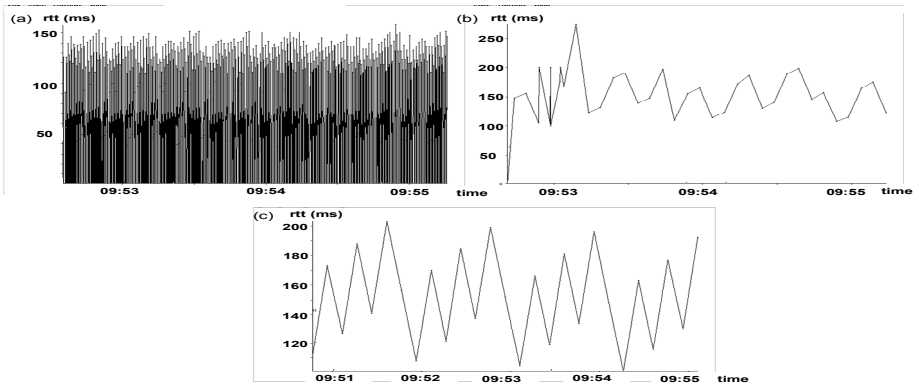

Figure 4. RTT measurement of PC-PLC sessions

Figure 4 illustrates the three sample RTT measurement graphs over the monitoring periods. The RTT values of each session are measured using tcptrace [10]. These PLC-PC sessions show clear periodic patterns which distinguish them from normal IP sessions. Their maximum RTT values also ranges from 150 ~ 250 msec, coinciding with the packet inter-arrival values of the sessions in Figure 3. The fixed packet arrival sequence is again apparent through the recurring shape of the graph.

It is important to recognize such pattern information of on-going sessions when detecting traffic anomaly and malfunctioning devices. This can be used as a guideline to determine 'irregularity' from the previously known communication patterns and avoid the ambiguous definition of anomaly in control IP networks.

3.3 Traffic Symmetry

Figure 5 shows the symmetric behavior in terms of number of packets being exchanged in a session. The upper plane (above 0) of the graph indicates packets per second (PPS) counts over the monitoring period in one direction. The lower plane shows the corresponding packet counts in reverse direction. All the graphs show almost identical packet transmission symmetry where one side shows slightly more packet counts than the other. These periodic packet generation patterns imply a simple request-response behavior between PC and PLC which follows a similar trail of HTTP behavior 9. However, unlike the HTTP traffic, the request object (or service) and the corresponding reply here are very much fixed in size and repetitive. The average PPS count of the sessions in control IP networks is below 10 PPS which is quite low compared to the Internet traffic.

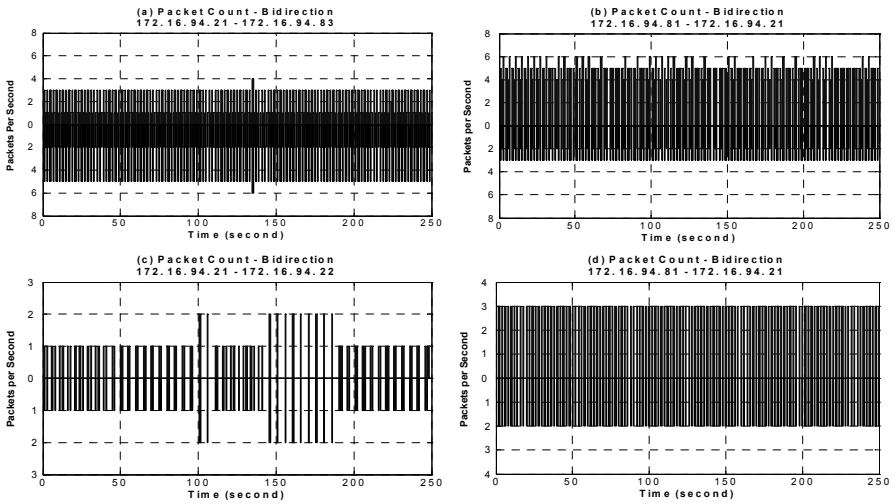

Figure 5. Bidirectional packets per second counts in PC-PLC sessions

3.4 Packet Size Distribution

The purpose of transmitting packets in control IP networks can be classified into the following: Signaling purpose for PLC operations, HMI display purpose, and management purpose (e.g., SNMP). At the time of data collection, we were assured that no management traffic was injected into the network. Figure 6 (a) and (c)

indicates the packet size distribution of packets mainly for signaling purpose. They show that over 90% packets are less than 100 bytes. In fact, their size ranges from 60 to 80 bytes which is just sufficient enough to contain TCP header information and a very few bytes for the actual payload. It explains for low bandwidth consumption in the control IP networks.

Figure 6 (b) shows the increase of average packet size compared to Figure 6 (a) and (c) where they are collected at near the edge PLC segment. At Segment B, we collect the HMI request/reply packets along with the PLC packets. The packet size distribution shows two distinguishable characteristics depending on where we collect the data – backbone or edge.

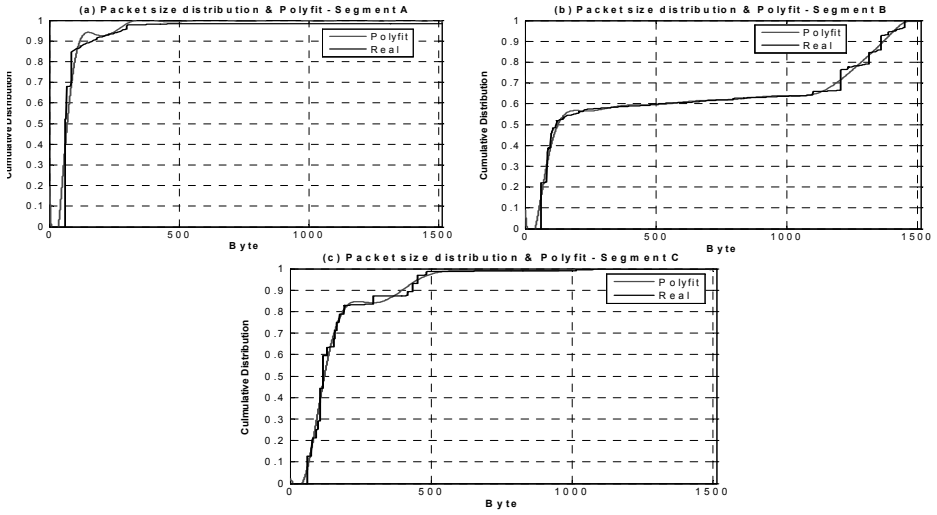

Figure 6. Packet size distribution – Segment A, B & C

3.5 Session Length Distribution

There exist two concrete patterns in session length: Short session length at the backbone (PC) and long session length near at the end host (PLC). Figure 7 (a) illustrates the session length distribution at Segment B. More than 90% of the total flows terminates less than 10 seconds. Note that, the active timeout value for flow export is set to 300 seconds which is determined by trail and error. This implies that the number of short periodical sessions outnumbers that of the long continuous sessions. The communications with the centralized PC servers are periodical, so that the flows from the identical hosts appears to be a separate flow. The communication period is relatively long, at least more than 300 seconds, for some PLC devices. At the edge of the network there exists more of continuous PLC communications. They are mixture of PC-PLC and PLC-PLC sessions in which the PLC-PLC sessions often stay within the local segment. Figure 7 (b) indicates the session tends to stay connected during the entire monitoring period, about 250 seconds. The session length distribution from another edge, Segment C, shows relatively longer session length than those in the backbone. However, it shows slightly uniform distribution of session length since the transmission intervals may vary due to different type of process.

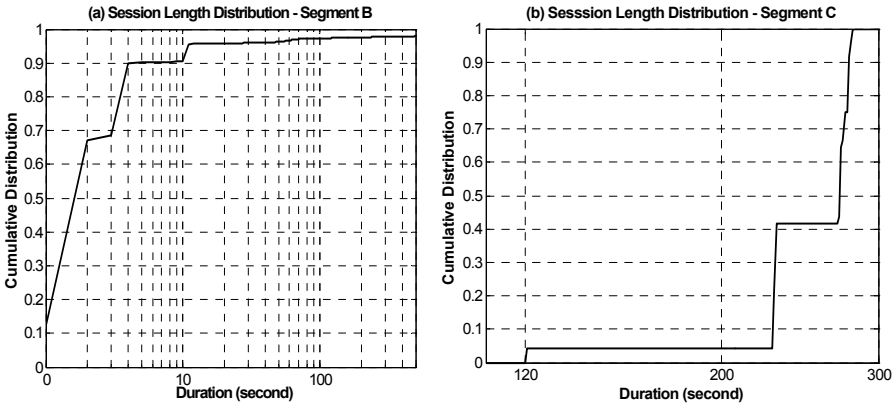

Figure 7. Session length distribution - Segment B & C

It is generally believed that the longer and continuous sessions occupy the control IP network because its operations often involve continuous and repetitive tasks over long hours (e.g., production line built on conveyor belt.) Overall, it follows a typical action-trigger behavior; we observe lighter commands (e.g., small packets) from the top of the hierarchy and corresponding actions at the bottom which triggers a complex sequence of communications.

3.6 Packet Reordering

We measure the following two categories to detect any packet reordering in a session: Out-of-sequence packets and retransmission packets. It refers to any out-of-order packet delivery. Figure 8 illustrates the ratio of the flows experiencing the out-of-sequence and retransmission packets. In the Segment B traces, we observe the total 21,539 TCP flows, and 92% of them experience one or more retransmission packets while online. The backbone traces have shown the sign of improper packet delivery compared to the rest of the traces. Despite of the high retransmission ratio, the process networks operates without any problem at the time of data collection. For future work, it is worthwhile to investigate whether such characteristic is bound to this particular case.

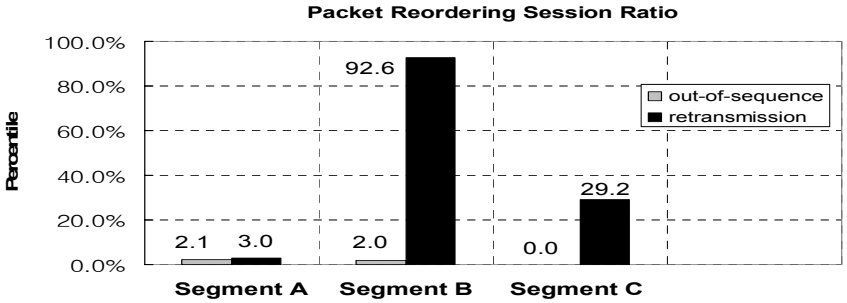

Figure 8. TCP sessions experiencing out-of-sequence and retransmission

4 Concluding Remarks

Despite of its large deployment in the real world, the area of monitoring and analyzing process control IP networks has not yet been focused very much by the network measurement and management research community. We have known very little details about the traffic behaviors of process control IP networks. These networks are deployed in mission critical operations which entail a maximum level of network stability. Understanding the traffic behavior helps us to operate fault-tolerant control IP networks where the cost of network malfunctioning is far more severe than the IP data networks.

In this paper, we have presented a measurement study of the traffic traces from the real industrial process control IP networks. The traces for analysis are carefully selected to provide a precise snapshot of the network from different perspectives, the traces from the backbone, the edge network, and the end host. We have summarized the following unique characteristics.

- Low-yielding and steady bandwidth usage regardless of monitoring periods
- Periodic traffic cycle in terms of packet arrival sequence and inter-arrival time
- Traffic symmetry
- Occurrence of small signaling packets
- Session length distribtuion patterns
- High packet reordering ratios at the backbone

Based on the primilnary analysis of traffic characteristics, we plan to develop a control IP network traffic model which can be used for network planning and anomaly detection. The comparison study between the process control networks and the ordinary IP networks will be also helpful to identify and formulate the systematical differences of process control networks.

References

1. Nobuo Okabe. “Issues of Control Networks When Introducing IP,” Proc. of Symposium on Applications and the Internet Workshops, Vol. 00, pp. 80-83, 2005.
2. Feng-Li Lian, James R. Moyne, and Dawn M. Tilbury. “Performance Evaluation of control networks: Ethernet, ControlNet, and DeviceNet,” IEEE Control System Magazine, 117(6), pp. 641-647, 2001.
3. Fieldbus Foundation. FF-581-1.3, “FOUNDATION Specification: System Architecture,” 2003.
4. PROFIBUS International. IEC 61158, “Digital Data communication for Measurement and Control – Fieldbus for Use in Industrial Control Systems,” 1999.
5. MODBUS.ORG, “Modbus Application Protocol V1.0,” 2002.
6. ASHRAE. ANSI/ASHRAE Standard 135-1995, “BACnet A Data Communication Protocol for Building Automation and Control Networks,” 1995.
7. EIA. EIA/CEA-709.1-B, “Control Network Protocol Specification,” 2002.

8. Libpcap, http://www.tcpdump.org/.
9. T. Kunz, T. Barry, X. Zhou, J.P. Black, and H.M. Mahoney. “WAP Traffic: Description and Comparison to WWW Traffic,” ACM Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems, August 2000.
10. Tcptrace, http://jarok.cs.ohiou.edu/software/tcptrace/.
11. POSCO, http://www.posco.com.