Actuator Quality in the Internet of Things

Julio Aráuz, *Member, IEEE* and Tony Fynn-Cudjoe.
J. Warren McClure School of Information and Telecommunication Systems,
Ohio University, Athens, OH 45701 USA

***Abstract*—In this paper we look at a generally unexplored area, the operational boundaries that determine perceived quality in actuators connected to the Internet of Things (IoT). Research in the area of IoT has placed emphasis on networked devices with sensing capabilities generally overlooking other types of applications. We believe that as the IoT grows it is important to understand what impacts the perceived quality of experience when a user commands networked actuators. Our research finds direct applicability in the control of cyberphysical systems such as those present in networked unmanned vehicles, autonomous control of automobiles or real-time industrial process control, among other options. Using a real packet-based actuating testbed along with qualitative and quantitative experimental data we study the quality experienced by actual users. We quantify how different network metrics affect perceived quality and propose the first characterization framework in the area.**

***Index Terms*—Internet of Things, cyberphysical systems, modeling**

I. Introduction

TRADITIONALLY the Internet of Things (IoT) has been considered as a set of interconnected devices with mainly sensing capabilities. For instance, the International Telecommunications Union views the IoT as a network of ubiquitous sensing devices with radio frequency identification tag handling functionalities. Furthermore, the ITU has ongoing efforts to create a service layer for machine-to-machine (M2M) communications under a global IoT set of sensing oriented specifications [1]. We believe it is important to understand not just sensing oriented applications of the IoT, but also applications of actuating networks. These are networks formed by actuators that carry out a physical task. Current research efforts in this area are scarce. We believe there are important and relevant future applications that demand an understanding of the quality of experience and operational limits under which networked actuating instruments may operate, especially when connected to network links.

Manuscript submitted May 02, 2013. Corresponding author: J. Aráuz (email: arauz@ohio.edu)

An actuator is a device that executes functions that involve movement to achieve a goal. Its actions can be carried out by using motors, electromagnetical elements, pneumatic or hydraulic systems, among other options. A networked actuator is a cyberphysical (CPS) system where the physical element involves mechanical movement. We will use the terms CPS or networked actuator interchangeably. Our work lies within an expanded vision of the IoT and looks at the basic implications of interconnected actuating objects handling a real time process. In such a case the perception of quality of the behavior of the system is critical as some tasks may require certain guarantees to operate correctly. For instance, certain tasks may only be carried out concurrently rather than sequentially, requiring guarantees that potentially numerous actuators, on possibly different types of networks, operate within certain network metric boundaries.

Examples of CPSs operating in real-time environments are the control systems of unmanned aerial vehicles, unmanned underwater near surface vehicles, actuators in experimental connected automobiles, among numerous other applications. However, these applications generally either assume a reliable communications channel is available, or are based on circuit switched control platforms. Research in that area has focused on measuring and improving link performance mainly on the physical not the cyber properties of the system [2][3]. In our case we explore the effects on qualitative and quantitative quality perception when controlling a real actuator over a packet switched unreliable link. In future applications these kinds of interconnected devices could be employed to control, in a distributed manner, real-time processes over links with varying quality.

In the next section we will start with a discussion of the motivation behind our study and then review related research in characterization of perceived quality of experience of real-time applications running over packet switched networks. Thereafter we will proceed to describe our experimental testbed and experiment design. We center our efforts on using experimental data to construct a first model for quality of experience of a networked actuator. We also provide a view of the operational boundaries where quality changes in network actuators, a critical factor for future applications in the area. Finally, we close with a discussion of results and our view of new research needs and opportunities.

II. Motivation

We believe that currently there are no formal characterization methods of the perceived quality of network interconnected actuators. We regard this field as one of great potential in numerous future applications drawing concepts from cyberphysical systems, human perception and networking. We envision these actuating systems performing real-time tasks that involve coordination among multiple entities. For instance, current experimentation efforts in autonomous road vehicles are looking at applications where one machine (e.g. controller) commands automobiles through wireless links in parking lots. Another growing field is that of unmanned aerial vehicles where devices can also be autonomously commanded over communication links. Future applications of networked actuators might also solve real-time needs in manufacturing, operation or maintenance in industrial settings.

There are a number of interesting challenges lying ahead in this field.

1) Network actuators will be expected to aid, perform or supervise tasks that traditionally require human intervention. However, any real-time application that involves human interaction and needs to be automated requires a quantifiable quality baseline. As an example, take the historical process by which analog voice quality over the public switched telephone network was evaluated. In that system a human based mean opinion score (MOS) was measured under strict experimental settings. This was done as it was not practical or possible to attempt, without a preexisting qualitative opinion, an accurate match between a quantitative characteristic from a signal (e.g. signal to noise ratio) and voice quality. With the advent of digital communications in the telephone network, again it was not feasible to attempt effective matching bit error ratio, signal to noise ratio or packet loss to voice quality without users' qualitative opinion. In general, processes that complement or enhance our human capabilities, like transmission of voice over a large distance, can directly benefit from a solid qualitative baseline. Furthermore, as discussed later, a baseline can be used to create a quantitative model of quality. In the VoIP domain carriers use such a model to autonomously monitor call quality by looking at basic network metrics.
2) Networked actuators will be expected to achieve multiple simultaneously goals in a continuous time environment. This is in conflict with the widely popular discrete time, sequential control systems we have today. The physical world requires concurrency while control systems semantics are largely sequential. This is a separate field from that of our interest in this paper. A thorough introduction to the joint dynamics and particular modeling techniques proposed in this field can be found in [4].
3) Quality of service guarantees for networked actuators that handle critical tasks (e.g. dangerous industrial processes or life related control systems) can benefit from the ability of the network to handle real-time distributed requests despite of challenging network conditions. Currently, M2M systems as well as other wireless systems like LTE or WiFi are already suited to handle real-time traffic under known formal restrictions (e.g. scheduler design or network load). However, there is still no formal understanding of the benefits and limitations of middleware distributed software that may complement a physical process; in the same way a basic dejitter buffer at any VoIP device is used to alleviate the effects of packet delay variation.

We focus our interest in this initial study on the first challenge. In this paper, we study a controlled actuating system to create a baseline and explore the impact of different network factors on perceived user quality. To the best of our knowledge this particular field has not been researched in the past in the domain of the IoT.

III. Related Work

We believe, there is no previous research that looks at how to directly estimate quality of experience for connected actuators over links with varying quality. However, there are numerous studies that have related goals in very distinct real-time domains. We will review the related research in four of these domains. First, we explore a successful approach to quantify and predict quality in VoIP systems. Thereafter we look at the characterization of quality in online games and take a look at quality of experience in video streaming applications. We close the section with a discussion on related efforts in the domain of human-computer interaction.

A. Predictive opinion score based on the R-Factor for VoIP

The R-factor [5] is a quantitative characterization of the performance of a VoIP application over a transport network. Its goal is to use transport level metrics and map them to perceived conversation quality. It enhances the ITU-T's E-Model which is used to map conversation quality to network design and equipment. To generate an opinion score this method requires channel quality metrics, along with pre-known architectural choices of coder implementation. Once the R-factor is calculated it can be directly mapped to a mean opinion score.

Several steps are required to construct the predictive opinion model used in the R-factor. First, transport level metrics data like delay, jitter and packet losses need to be collected. Then, using these transport metrics along with de-jitter buffer information, packet size and the coder frame data, a structure known as an *error mask* is constructed. The error mask is essentially a description of the sequence of good and bad frames. Thereafter a lookup table, based on qualitative metrics, is used to map the results from the mask to the E-Model's equipment impairment. The equipment impairment is the key value needed, along with the network metrics, to generate the actual R-factor value.

The E-Model originally characterizes the R-factor as the sum of four terms [6].

$$R = 100 - I_s - I_d - I_{ef} + A \tag{1}$$

Where I_s is the impairment due to signal-to-noise level, I_d is the impairment due to delay in the path, I_{ef} characterizes losses within gateway codecs and A is called the expectation factor. The value of A tries to measure the subjective quality appreciation of users that under some conditions (e.g. mobile phone conversations, or lower prices) are willing to grade a conversation as acceptable.

Equation (1) can be further simplified with experimental data and by disregarding the effects of the expectation factor. Resulting in $R = 94.2 - I_d - I_{ef}$. To further simplify the expression, a piecewise linear approximation of the value of I_d as a function of the one way mouth-to-ear delay was constructed by the authors; leaving I_{ef} as the only element needing formal characterization. This last parameter is derived from subjective measurements for distinct codecs and the different network metrics mentioned above and can be mapped to packet loss information. In [5] the authors then show how the R-factor can be expressed as a function of delay and packet loss. The delay term is in turn a function of the delays introduced by the codec, the de-jitter buffer and the network itself. All of which can be directly accessible from measurements adding to the practicality of the R-factor for automated conversation quality monitoring.

B. Quality in networked games, the G-Model

In a different domain, online gaming, an area of continuous growth, quality prediction has also been formally looked at. In [7] the authors constructed a quality prediction model based on end-to-end network quality measurements. As in the case of the R-factor, where data is available for just a subset of configurations, this study also focuses on just one type of game (i.e. first person shooter Quake IV) and on measuring just one specific goal inside the game.

On the network metric side, the authors considered the end-to-end roundtrip time, jitter and packet loss as factors to quantitatively characterize quality. In this case the different levels these factors assume, impact a player's ability to carry out specific tasks in a gaming environment. In particular, in [7] the authors tested the participants' ability to perform a task under varying network conditions asking them to subjectively judge quality for distinct setups using the absolute category rating suggested by the ITU-T for voice applications [8]. This scale incorporates the widely used categories for qualitative scoring (i.e. excellent = 5, good = 4, fair = 3, poor = 2, bad = 1).

For their specific gaming application the authors found significant degradation of the mean opinion score when end-to-end roundtrip time was above 80 ms. A similar degradation started at 40 ms of average jitter. Interestingly the authors did not find any noticeable qualitative degradation when packet losses where introduced; however they did not explore the reasons. We speculate that a likely cause is the presence of a reliable data transfer protocol (likely TCP) that allows fast packet recovery before the end user notices any degradation; however further confirmation evidence is needed.

The authors used a standard multivariate regression analysis along with their experimental data to construct a prediction model, the G-model, of the compound network metrics effects on perceived user quality. The analysis found that the regression model used provided a good match between the qualitative data and the prediction.

C. Perceived quality in video streaming

Another real-time domain that requires both qualitative and quantitative means to evaluate quality is video streaming. As the number of companies streaming stored or live video on the Internet grows, it is vital to count with means that quantify user behaviors (e.g. abandonment rate, quality rating, etc). A detailed causality study between viewer behavior and video streaming quality is presented in [9].

Their findings are quite interesting. Using a massive data set of online video viewing statistics the authors found that an average of two seconds of startup delay is generally enough for users to consider abandoning the idea of watching a video. Similarly, just a small percentage of buffering, 1% of the duration of a video, impacts by 5% the total viewing time a user dedicates to a video. These findings clearly illustrate similar relations to those we are interested in between perceived quality and basic network metrics.

To reach their conclusions the authors used both Kendalls correlation and a causal analysis known as quasi experimental design (QED). This causal analysis technique is needed as the authors identified several confounding variables (e.g. type of video content, connection type and viewer geographical origin) that may skew the results. In QED viewers with similar confounding variable values are matched together and only then causality is studied, thus reducing the effect of these variables.

D. Fitt's law

In the human-computer interaction domain an interesting approach to evaluate a user's capability to move a pointer on a computer display is known as Fitt's law. This law does not include network metrics but we include it here, as our testbed asks users to perform a similar task in a cyberphysical domain. In [10] the authors discuss both Fitt's law and several of its variations and extensions to multi-dimensional tasks. Interestingly the law has a formulation similar to that of Shannon's theorem for channel capacity. Equation (2) shows the formulation.

$$MT = a + b \log_2(1 + \frac{2A}{W}) \tag{2}$$

Where MT is the time to move a pointer to a target of width W located a distance A away. The values of a and b can be determined via regression analysis of experimental data and are functions of the speed of movement, start and stopping time as well as users' average abilities. We mention this approach due to the similarity of the task it characterizes to our approach. It provides useful insight for evaluating human-machine interactions, an aspect that could be employed later to characterize quality for CPSs as well.

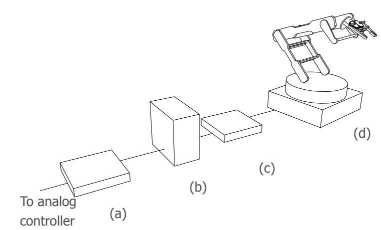

Fig. 1: Experimental testbed

E. Previous approaches in perspective

In our research we are interested in understanding the effects of different variables that may affect the perception of quality of a given set of tasks. Therefore our proposal shares some foundations with all the previous efforts discussed above. In particular, we will be looking at a similar set of network metrics and their effect on perceived quality. Our data comes from an experimental testbed and data collection through first-time user experimentation. In our case all experiments were carried out under the same conditions and types of users, thus we do not place emphasis on the effects of any confounding variables. As discussed next, the findings and methodologies used in these previous studies are related to our approach. They also provide a direction for future possibilities and enhance the understanding of perceived quality.

IV. Experimental Testbed

In light of the need to make our testbed as generic as possible we selected an actuating system with numerous degrees of freedom rather than a common servo-controlled linear actuator. Our testbed consists of a electro-mechanical arm controlled by direct current motors along with gear boxes that enable movement in several degrees of freedom. We consider, for this first study, up to two degrees of freedom. Using this testbed we obtain qualitative and quantitative data of the perceived quality of the experience of first-time users commanding the system to achieve a set of goals.

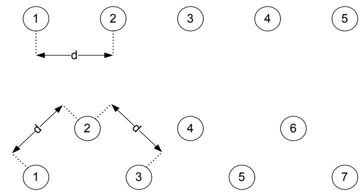

Fig. 2: Goal setup, distribution of targets. One degree of freedom (top). Two degrees of freedom (bottom).

A. Physical system

The actuating system, a mechanical arm, is controlled by a set of microcontroller boards as illustrated in Fig.1. The first board *(a)*, shown on the left side, is also attached to an analog set of mechanical controls for user operation. On the right side it is attached, via an Ethernet interface, to a link emulator *(b)*. The emulator introduces variations in the levels of network metrics such as delay and packet loss. The second board *(c)* is attached to the link emulator on one side and incorporates the necessary electronics (i.e. H bridges) to independently actuate two direct current motors on the other side. In the actuating system, the electromechanical arm *(d)*, each motor can rotate at constant speed in any direction allowing only horizontal or vertical plane movement.

The analog controller given to the user allows the command of the system via two single pole, double throw spring loaded switches. Each switch can place a motor in any of three states: clockwise rotation (cc), no movement *(off)*, or counterclockwise rotation (ccc). To activate a motor a user needs to hold the control in the desired position. To stop it the user just needs to release the control as the spring mechanism will move back the switch to the idle position.

Our goal is to measure users' perceived quality of experience when trying to achieve quantifiable goals. We achieved this by designing a set of experiments that required the user to either employ one degree of freedom (horizontal movement) of the arm or two degrees of freedom (horizontal and vertical movement). The experiment required users to direct a fixed laser attached to the fixed arm's grabber towards a set of targets located on the vertical plane a fixed distance from the arm.

This setup is illustrated in Fig. 2. For the first task, with targets illustrated in the top subfigure, users had to direct the laser towards a set of horizontal, equally spaced, ordered targets using only the horizontal movement control. Each target consisted of a circle with a ¾ inch diameter. The second task, illustrated in the bottom subfigure, demanded users to use simultaneously both horizontal and vertical movement controls to point the laser to targets equally spaced in two dimensions, again located on the vertical plane. For both tasks the distance d was fixed and equal to 25 cm and goals were considered achieved when the laser created point of light was fully inside a target.

B. Packet based control system

The actuating system motors are activated using a packet based system. We employ UDP as the transport protocol over the links to avoid including any confounding effects from retransmissions from protocols such as TCP. We decided to use a *one packet state transition mechanism* to command the motors. In this approach the payload carried by a packet specifies a state φ among the set of states $\zeta =${*cc, off, ccc*} and the motor index i, ($i \in \{1, 2\}$). When a user places a control switch in a movement state, one packet is sent from the first microcontroller board through the link emulator to the second board, specifying the duple $\{\varphi, i\}$. One corresponding packet specifying to stop movement is sent once the user releases the command switch. Thus any motor changes state after the successful reception of one packet. This differs from an approach we do not explore where the state of the switches is continuously sampled and packets sent continuously as well.

V. Evaluation Methodology

A. Experimental Design

Our goal is to match the qualitative user opinions against quantitative measured performance. Thereafter, characterize a model to predict quality of experience based on multiple factors. To do this we asked each user to rate the quality of experience for each experiment using the scale detailed in section III-B. We used as experimental factors: average one way delay, packet loss, and the number of degrees of freedom.

The average delay factor, γ, corresponds to experiments carried out with one-way delay values drawn from discrete uniform distributions with standard deviation σ. Since our goal was to consistently measure the quality of experience at each level, we made $\sigma << \alpha$. Where α is the observed maximum deviation from a mean before a user typically changes his opinion score of the system. Then, based on early tests, we chose $\sigma = (1/3)\alpha$, which resulted in a σ value of 30 ms. We experimented with five levels of γ (ms) ($\gamma \in \{0, 100, 200, 300, 400\}$). Additionally we took into consideration five levels for the packet loss probability, $\rho(\%)$, ($\rho \in \{1, 5, 10, 15, 20\}$). Finally,

TABLE I: Tabulated results for all combinations of factors and levels for one degree of freedom

		One Degree of Freedom ($\tau = 1$)			
		Completion Time		Opinion Score	
Factor	Level	Mean (s)	95% C.I.*	Mean	95% C.I.
Average delay γ (ms)	0	2.30	(2.17, 2.44)	4.13	(3.98, 4.28)
	100	2.90	(2.74, 3.07)	4.38	(4.17, 4.60)
	200	3.40	(3.15, 3.66)	3.88	(3.73, 4.03)
	300	3.93	(3.63, 4.24)	3.13	(2.86, 3.40)
	400	4.5	(3.82, 5.19)	2.5	(2.01, 2.99)
Packet loss ρ (%)	0	2.30	(2.17, 2.44)	4.25	(4.06, 4.45)
	5	2.73	(2.65, 2.82)	4.00	(3.78, 4.23)
	10	3.44	(3.18, 3.71)	3.25	(2.82, 3.69)
	15	4.64	(4.21, 5.07)	2.50	(2.19, 2.82)
	20	5.02	(3.86, 6.19)	2	-

* C.I. = confidence interval

TABLE II: Tabulated results for all combinations of factors and levels for two degrees of freedom

		Two Degrees of Freedom ($\tau = 2$)			
		Completion Time		Opinion Score	
Factor	Level	Mean (s)	95% C.I.*	Mean	95% C.I.
Average delay γ (ms)	0	5.62	(5.17, 6.08)	3.75	(3.46, 4.05)
	100	8.67	(7.86, 9.49)	3.38	(3.17, 3.60)
	200	11.83	(10.68, 12.98)	2.63	(2.42, 2.85)
	300	18.49	(16.46, 20.53)	2.13	(1.72, 2.55)
	400	23.35	(18.40, 28.30)	2	-
Packet loss ρ (%)	0	5.62	(5.16, 6.08)	3.75	(3.46, 4.05)
	5	8.57	(7.70, 9.45)	3.00	(2.69, 3.32)
	10	6.47	(5.88, 7.07)	2.63	(2.25, 3.02)
	15	11.50	(10.26, 12.74)	1.88	(1.61, 2.15)
	20	13.64	(11.26, 16.02)	1.5	(1.01, 1.99)

* C.I. = confidence interval

for the number of degrees of freedom, τ, we set two levels, $\tau \in \{1, 2\}$. Corresponding to horizontal movement only ($\tau = 1$), and horizontal plus vertical movement ($\tau = 2$).

As we employ a *one packet* based control to evolve the state of the system we do not evaluate performance for individual levels of fixed delay (e.g. $\sigma = 0$) or jitter. We do not place emphasis on such a study as we found out that in general users can rather easily predict the arm's movement to achieve a goal when the delay is fixed. We do look at varying levels of average delay in the network, a condition that does replicate real use cases of a networked connected cyberphysical system.

Experimental data was acquired by giving the same set of experiments and goals to eight first-time users. Experiments were carried out under equal external conditions to minimize the effect of any environmental confounding variables (e.g. appropriate lighting, minimal distractions, etc.). For each combination of factors and levels, the time a user needed to move the point of light between targets was measured. Therefore, for each user and one degree of freedom targets (top of Fig. 2) four times were recorded. For two degree of freedom targets (bottom of Fig. 2) six results were collected per each combination of factors and levels. At the end of each set of tasks the user was asked to qualitatively categorize the experience using the opinion score scale [8].

VI. Analysis

Tables I and II show the results of carrying out separate experiments for different levels of γ, ρ and τ. We kept either γ or ρ at zero for the corresponding tests. The table includes the mean values for both the completion time (e.g. average time needed to move the point of light between two targets) and the opinion score along with their 95% confidence intervals. We will later use these results to verify suitability for a regression study.

A first analysis of the results yielded some interesting conclusions. In all cases, except for one, as the delay and packet loss conditions deteriorate the average time to complete a task increases and the qualitative opinion scores decreases. The variations in both outcomes are in general statistically significant especially for cases where there are good network conditions and thus high opinion scores. However when conditions deteriorate (e.g. $\gamma \geq 300$ ms and

$\rho \geq 15\%$) the results start to become the same as the confidence intervals tend to overlap including mean values inside each other. This happens as under challenging network conditions the standard deviation of the time to complete a task grows for all users. In regards to the number of degrees of freedom, τ. As expected the higher the degree the worst the quality experienced both in quantitative and qualitative terms. Furthermore, in two cases when (γ, ρ, τ) was equal to (400, 0%, 2) and (0, 20%, 1) all users agreed on the same opinion score for the system.

There is an interesting data point that required additional examination and verification, when ρ varies from 1% to 5% under two degrees of freedom. Here the average time to complete a task decreased which is unexpected. After careful data verification, we realized that under these particular conditions users were able to habituate to the experiment and gained experience in handling the system; resulting in lower average times. Interestingly, this was the only case where we observed this. Although this is a limitation of the testing methodology, it helped us gain valuable insight for future evaluation sets.

With these experimental results we want to gain insight of the system from two distinct perspectives:

1) The impact of each factor on the mean opinion score.
2) The ability to predict perceived quality based on the quantified impact of each factor.

We employed a factorial design statistical method to evaluate the impact of factors and a multiple linear regression analysis to predict estimated responses based on the data.

A. 2^k *Factorial Design Analysis*

In an ideal setup, data sets include a vast number of combinations of factors and levels. In most experiments obtaining such a set is either too expensive or requires a long period of observation. Factorial design is a technique that allows quantifying the impact of factors on an outcome by looking at just two levels per factor [11]. These levels are usually chosen from the data set from those that quantitatively result in statistically different outcomes.

We carried out a 2^k ($k = 3$ factors) factorial design analysis with the levels detailed in Table III. A subset of the combined effects data needed for the analysis can be obtained from Tables I and II. In particular, $\gamma = 0$, $\rho = 0$ can be obtained from the first row of each table; $\gamma = 0$, $\rho = 20\%$ from the last one. For the rest of the combinations we carried out separate experiments that indicated that, under the worst conditions of average delay and packet losses from Table III, the opinion score consistently dropped to 1.0.

TABLE III: Factors and Levels for 2^k Design ($k = 3$)

Factor	Level -1	Level 1
Average delay γ in ms	0	400
Packet loss ρ	0%	20%
Degrees of freedom τ	1	2

The 2^k design enables an understanding of the factor impact, as a nonlinear regression function of the factors. In formal terms, let $\Gamma = P = T = -1$ be variables corresponding to factors γ, ρ and τ respectively for Level -1. Similarly, $\Gamma = P = T = 1$ for Level 1. Then M, the mean opinion score, can be expressed as:

$$\begin{aligned} M \quad = \quad & q_0 + q_\gamma \Gamma + q_\rho P + q_C T + q_{\gamma\rho} \Gamma P + \\ & q_{\gamma\tau} \Gamma T + q_{\rho\tau} PT + q_{\gamma\rho\tau} \Gamma PT \end{aligned} \tag{3}$$

Where q_0 represents the mean of M. The effect of each respective factor and their combined effect on the mean opinion is represented by q_i, ($i \in \{\gamma, \rho, \tau, \gamma\rho, \gamma\tau, \rho\tau, \gamma\rho\tau\}$). Table IV shows the factorial design results for all q_i along withe their respective variations.

TABLE IV: Factorial Analysis Results

Parameter	Mean Estimate	Variation Explained (%)
q_0	2.2350	-
q_γ	-0.6100	30.82
q_ρ	-0.8600	61.27
q_τ	-0.1725	2.46
$q_{\gamma\rho}$	0.2350	4.57
$q_{\gamma\tau}$	0.0475	0.19
$q_{\rho\tau}$	0.0475	0.19
$q_{\gamma\rho\tau}$	0.0775	0.5

The results are interesting and shed light on the impact of the different factors. The analysis yields a mean opinion score, M of 2.2350. M is mostly affected by packet losses which makes a difference of $\pm$0.8600 and explains 61.27% of its variation. The average delay makes a difference of $\pm$0.6100 and explains 30.82% of the variation. The number of degrees of freedom makes a difference of $\pm$0.1725 and contributes only 2.46% of the variation. The effects of the combination of the other factors make up a contribution of only 5.45%.

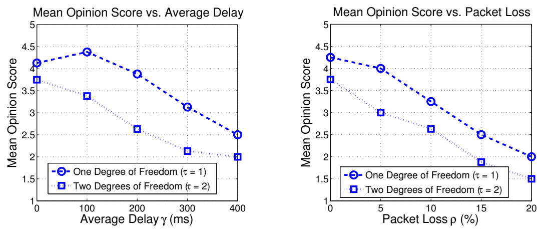

Fig. 3: Mean opinion score results vs. γ (left), and vs. ρ (right) (Conf. Interv. shown in Tables I and II)

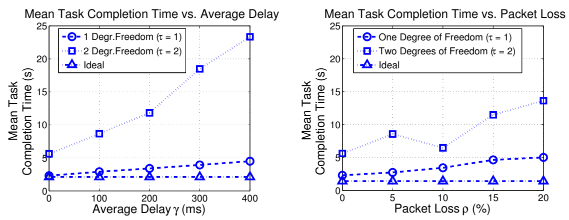

Fig. 4: Mean task completion time vs. γ (left), and vs. ρ (right) (Conf. Interv. shown in Tables I and II)

B. Multiple Linear Regression

To predict the value of the mean opinion score, M, as a function of the factors and their levels we employed a multiple linear regression analysis. As a preamble to the study we looked at the relationship between each independent variable (e.g. factors) and the dependent variable M. Figure 3 shows these dependencies for the effects of each factor on the output for both degrees of freedom. The visual inspection of the trends suggests linear relationships between M and γ, and M and ρ.

Therefore, it is possible to express the predicted value of the mean opinion score, M_i as a function of the dominant factors determined in the factorial analysis, as shown in Eq. (4).

$$M_i = b_0 + b_1\gamma_i + b_2\rho_i \tag{4}$$

Where γ_i should be expressed in milliseconds and ρ_i as a numerical probability fraction, and b_i ($i \in \{0, 1, 2\}$) are the parameters of the regression. The results are shown in Table V. These indicate that all the regression parameters had significance at a 95% confidence level; as they do not include zero in their confidence interval. Separately, we also visually verified independence of errors and did a quantile-quantile check for normal distribution of errors.

Regarding the computed coefficient of determination, it indicates that 92% in the variation of M can be explained by the regression. This value, although not ideal (e.g. closer to 100%), indicates that the regression is a good model for the range of values tested. A basic multicollinearity analysis was also carried out by computing the correlation among predictors, which indicated little to no correlation. Thus collinearity does not exists and both predictors should be used in the model.

TABLE V: Multilinear Regression Results

Parameter	Value	95% C.I.
b_0	3.71	(3.09, 4.32)
b_1	-0.00305	(-0.0048, -0.0013)
b_2	- 8.60	(-12.15, -5.05)
Coeff. of determination R^2	0.92	
Correl. among predictors	0.01	
Std. deviation of errors	0.39	
F statistic	29.10	
F tabulated	5.79	

We also carried out an analysis of variance. The results indicated that the probability of the F statistic computed value being obtained by chance is just 0.17%. As a first set of results we believe the regression provides useful impact on the behavior of the testbed, although further data to improve its effectiveness in terms of R^2 and the F statistic will be quite beneficial.

C. Task Completion Analysis

As discussed earlier, tests were carried out for independently testing one factor while keeping the other one at a zero level. These correspond to the results from Tables I and II and plotted in Fig. 4. The trends clearly show the monotone impact of average delay on task completion time. Regarding packet losses, we already discussed the apparently out of ordinary impact of packet loss earlier in this section.

It is feasible to compute the ideal time to perform a task in any of our configurations as the constant velocity at which the system can move in any direction and the distance between the center of the targets is known. This is also illustrated in Fig. 4. This ideal result provides a baseline to quantitatively understand the mean effectiveness of a user to control the actuators in relation to a perfect no delay, no losses condition. In an ideal setup two seconds are required to go from one target to the next, for one degree of freedom, and 1.42 seconds in a two degree of freedom configuration. Notice that in the worst cases for (γ, ρ, τ) in each figure, there are differences of approximately 21 seconds and 12 seconds in task completion time in relation to the ideal scenario.

This illustrates the difficulty imposed by the network impairments, difficulty that is then also observed in the opinion score.

VII. Conclusions and Future Work

Following similar approaches to those present in other domains we constructed a statistically accurate first quality model for a basic system of actuators. We found that the main factors that impact perceived quality were average delay and packet loss, not the number of degrees of freedom. Delay taken independently degraded quality once it reached an average value above 300 ms. Packet losses of above 10% also resulted in heavily degraded quality. This is mainly due to the nature of the system which changes state with just one packet, therefore if a packet is lost it is rather easy for a user to completely miss a target.

The factorial analysis showed packet loss as the dominant factor affecting perceived quality. Furthermore, the results provided an interesting insight on the magnitude of the effect of the packet loss which is twice that of delay. This has some contrast to what is experienced by VoIP applications where less than a percentage point of packet loss is considered viable before quality degrades; however call quality stays relatively high over a range of average packet delay values.

Our results open the door for an understanding of other interesting conditions. For instance, impact on quality with different packet control strategies. More importantly, how to construct a middleware that can mitigate network effects and give consistent, or perhaps adaptive usage quality, to the user. As the IoT grows in applications and services such a middleware will require results from further experimentations with real testbeds such as the one proposed in this paper.

References

[1] "ITU-T Global Standards for the Internet of Things," http://www.itu.int/en/ITU-T/techwatch/Pages/internetofthings.aspx, Accessed: March 01, 2013.

[2] E. Pfeifer and F. Kassab, "Dynamic feedback controller of an unmanned aerial vehicle," in *Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), 2012 Brazilian*, Oct., pp. 261–266.

[3] P. Zhan, K. Yu, and A. Swindlehurst, "Wireless relay communications with unmanned aerial vehicles: Performance and optimization," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 47, no. 3, pp. 2068–2085, July 2011.

[4] P. Derler, E. Lee, and A.-S. Vincentelli, "Modeling Cyber Physical Systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 13–28, 2012.

[5] R. G. Cole and J. H. Rosenbluth, "Voice Over IP Performance Monitoring," *SIGCOMM Comput. Commun. Rev.*, vol. 31, no. 2, pp. 9–24, Apr. 2001. [Online]. Available: http://doi.acm.org/10.1145/505666.505669

[6] ITU-T, "Recommendation G.107, The E-Model a Computational Model for use in Transmission Planning," December 1998.

[7] A. F. Wattimena, R. E. Kooij, J. M. van Vugt, and O. K. Ahmed, "Predicting the perceived quality of a first person shooter: the Quake IV G-Model," in *Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games*, ser. NetGames '06. New York, NY, USA: ACM, 2006. [Online]. Available: http://doi.acm.org/10.1145/1230040.1230052

[8] ITU-T, "Recommendation P.800, Methods for subjective determination of transmission quality," August 1996.

[9] S. S. Krishnan and R. K. Sitaraman, "Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs," in *Proceedings of the 2012 ACM conference on Internet measurement conference*, ser. IMC '12. New York, NY, USA: ACM, 2012, pp. 211–224. [Online]. Available: http://doi.acm.org/10.1145/2398776.2398799

[10] I. S. MacKenzie and W. Buxton, "Extending Fitts' law to two-dimensional tasks," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '92. New York, NY, USA: ACM, 1992, pp. 219–226. [Online]. Available: http://doi.acm.org/10.1145/142750.142794

[11] R. Jain, *The Art of Computer Systems Performance Analysis*. John Wiley & Sons, New York, 1991.

Julio Aráuz is an Assistant Professor at the J. Warren McClure School of Information and Telecommunications Systems at Ohio University. He obtained both a Masters in Telecommunications and a Ph.D degree in Information Sciences from the University of Pittsburgh. Before joining Ohio University he was a Research Scientist at NEC Network Laboratories in Heidelberg, Germany working for the Mobile Wireless Network Division. His research interests include wireless communications, cyberphysical and self-organizing complex systems, simulation and testbed experimentation.

Tony Fynn-Cudjoe received his bachelor's degree in Computer Engineering from Kwame Nkrumah University of Science and Technology, Ghana. Currently, he is a pursuing a master's degree in Information and Telecommunication Systems at the J. Warren McClure School of Information and Telecommunications Systems at Ohio University. His research interests include computer networks and its reliability, wireless networks and system design. Prior to pursuing his master's degree, he held various positions as a network engineer for various ICT companies in Ghana.