

Project	IEEE 802.16 Broadband Wireless Access Working Group	
Title	Proposal Evaluation Form and Scoring Process [Draft]	
Date Submitted	1999-10-20	
Source	Roger Marks NIST 325 Broadway, MC 813.00 Boulder, CO 80303	Voice: +1-303-497-3037 Fax: +1-303-497-7828 E-mail: r.b.marks@ieee.org
Re:	IEEE 802.16-99/05 ("Development Plan for the 802.16.1 Air Interface Standard") and discussion on scoring systems at Closing Plenary of Session #3.	
Abstract	This document discusses issues involved in the 802.16 proposal evaluation and scoring process. It presents some guidelines for the process based on the issues involved. It suggests that individual voters should record their scores using a spreadsheet and submit them using a text. The best way to handle the submitted text files is not specified. The document also considers various possibilities for normalizing the scores and suggests a gentle approach using a cap on the average deviation of the scores from a nominal value. A sample spreadsheet form accompanies this document.	
Purpose	The document should lead to an effective proposal scoring process. Comments and suggested improvements are requested.	
Notice	This document has been prepared to assist the IEEE 802.16. It is offered as a basis for discussion and is not binding on the contributing individual(s) or organization(s). The material in this document is subject to change in form and content after further study. The contributor(s) reserve(s) the right to add, amend or withdraw material contained herein.	
Release	The contributor acknowledges and accepts that this contribution may be made public by 802.16.	
IEEE Patent Policy	The contributor is familiar with the IEEE Patent Policy, which is set forth in the IEEE-SA Standards Board Bylaws < http://standards.ieee.org/guides/bylaws > and includes the statement: "IEEE standards may include the known use of patent(s), including patent applications, if there is technical justification in the opinion of the standards-developing committee and provided the IEEE receives assurance from the patent holder that it will license applicants under reasonable terms and conditions for the purpose of implementing the standard."	

Proposal Evaluation Form and Scoring Process

Roger Marks
Chair, 802.16

Abstract

This document discusses issues involved in the 802.16 proposal evaluation and scoring process. It presents some guidelines for the process based on the issues involved. It suggests that individual voters should record their scores using a spreadsheet and submit them using a text. The best way to handle the submitted text files is not specified. The document also considers various possibilities for normalizing the scores and suggests a gentle approach using a cap on the average deviation of the scores from a nominal value. A sample spreadsheet form accompanies this document.

Introduction

Document IEEE 802.16-99/05 ("Development Plan for the 802.16.1 Air Interface Standard") describes a proposal evaluation process based on scoring by the voting members, in several categories, on a 0-10 basis. In the early stages, proposers scoring better than some minimum value in any category will be invited to resubmit for the following session; at later stages, a high score average across all categories is required. IEEE 802.16-99/05 refers to an "802.16 scoring procedure" which has not yet been defined. Such a procedure was discussed at 802.16 Session #3, including during the Closing Plenary, at which score normalization was debated.

Proposal Evaluation Form: Considerations and Solutions

- The MAC and PHY Call for Contributions (IEEE 802.16m-99/01 and IEEE 802.16p-99/01) each specify more than 10 evaluation criteria. This could require each voting member to submit hundreds of scores. This scoring must therefore be handled electronically.
- IEEE 802.16p-99/01 describes a process in which, at Session #4, the scores are cast at the end of the Thursday morning session. Invitations to submit for Session #5 should be finalized in the Closing Plenary on Thursday afternoon. This short time frame again mandates electronic submission and electronic score processing.
- Not all computers have floppy drives, PC card slots, or Ethernet ports. Therefore, we need flexibility in how the data is to be submitted. I suggest that the most flexible format is a file, since this can be submitted on a floppy, on a flash ROM card, or by Ethernet (if we can set up an Ethernet hub and server).
- Not all voting members will bring computers, so some sharing will be required. Again, flexibility in submission alternatives will ease the process.
- It is not clear how all of the scorers' data will be combined. To maintain flexibility, we need a file format that any program can read. I suggest that ASCII is the best choice.
- Some form of normalization or raw-score modification may be used. If so, it will be helpful to give the submitter an immediate display of the normalized or modified score.
- The form ought to run on a popular software application.
- Given these requirements, I believe that a Microsoft Excel spreadsheet will be a good application to host the scoring form. The data can be easily exported into a text file.
- The form should be distributed in advance of Session #4 so that voting members can begin initial scoring of proposals based on the submitted documents, which are due October 29.

Normalization: Considerations and Solutions

- When the scores of individual voters are combined, the most sensible approach is to use a simple average. All voters should be considered equally. For instance, we must not attempt to reduce the ability of members to evaluate proposals in which they have a particular interest.
- If any normalization is to be done, it should be based on the scoring data provided by each voter independently. That way, each voter can know the effective vote that accompanies an actual vote.
- At Session #3, Mr. Marin that suggested that each voter's scores, in each category, be normalized to the some nominal mean (possibly 5 out of 10) and nominal standard deviation. Objections were raised to the normalized mean on the grounds that this would in effect turn the scoring system into a ranking system, counter to the intent of the IEEE 802.16-99/05. For instance, the system would force about half of the scores below the nominal mean, effectively forcing them to fail to meet the minimum score for further consideration.
- On the other hand, there is a clear motivation for providing some kind of normalization to the deviations of the scores. Namely, many voters are not neutral judges but will have distinct biases. As a result, some participants may be driven to maximize the effect of their vote by scoring their favorite proposal as a 10 and all others as a 0. This will not contribute positively to the consensus process and will not optimize the value of the evaluation process as a feedback mechanism to the proposers. Furthermore, it contradicts the 802 rules (the document "Operating Rules of IEEE Project 802 LAN MAN Standards Committee"), which specify that "Working Group members shall participate in the consensus process in a manner consistent with their professional expert opinion as individuals, and not as organizational representatives". If we can use the scoring rules to limit the tendency toward extreme voting, we should end up with more useful results.
- It is possible to renormalize all votes to the same standard deviation. However, this doesn't really address the problem. For a large set of proposals, the standard deviation of a set of scores including all zeroes except for one ten does not yield an extremely high standard deviation because the mean is close to zero.
- It can be assumed that most voters do not appreciate having their scores externally revised or renormalized. Therefore, a good scoring restriction system places limits; voters operating inside those limits will have their scores recorded exactly as given.

With these issues in mind, I am suggesting a scoring restriction based on the idea that the "nominal" score is 5 and the normal range of scores is 2.5-7.5. Scores outside of this range are allowed as the exception but not the rule. The algorithm works as follows:

- Compute the absolute value of the deviation of the each score (by a single voter in a single voting category) from the nominal score of 5. Average these for an average deviation.
- If that average deviation is less than 2.5, I leave the voter's scores alone.
- If the average deviation is over 2.5, scale the scores linearly back towards 5 until the average deviation is 2.5. This effectively caps the average deviation at 2.5.

This system should leave most scores unretouched, but people will be prohibited from scoring everything out near the edges of the scale. It is intended as a gentle control but one that restrains voters from going to extremes.

Draft Scoring Form

File 802.16-99/07a.xls is a sample Excel spreadsheet that demonstrates how the data can be easily recorded, restricted by capping the average deviation (with the caps indicated to the scorer), and exported to a text file. Two of these forms would be needed, one each for the MAC and PHY evaluations.

Comments and Suggestions

I solicit your comments and suggestions on the spreadsheet, on the average deviation cap, on ideas for collecting and processing the data of many evaluators, and on better ideas for handling the evaluation process.