The Ins and Outs of Scouting in FIRST Robotics Competitions

By Joseph Lewis, FIRST Team 701: The RoboVikes

Introduction

This is designed as a guide to future and current Scouting Directors to better understand how to scout FRC Events. This guide is intended for use by any team looking at a better way of organizing their current scouting system, starting a new scouting system, or even simply wanting a fresh perspective on the matter. Sections of this guide detail the current scouting system that FIRST Team 701: The RoboVikes uses at FRC Events to further their competitive gameplay.

This in no way is presented as the only proper way of scouting, nor the best or necessary way. It is simply an outline of the practices that FIRST Team 701: The RoboVikes has found that work the best for them. The guide really works best if you read the entire document.

Scouting Defined

Scouting is roughly defined as a process in which information about teams at an event is both **collected, processed,** and **interpreted.** Its purpose is simple - to further the competitive ability of the team collecting and processing that information. Its implementation, however, is very complex. There are many methods to go about collecting, processing, and interpreting data. Collecting the right information and processing it quickly and efficiently and interpreting it correctly can mean the difference between winning and losing not only a match, but also an entire event.

Collecting Information

Types of Information

The first part of scouting relies on collecting information about the teams at an event. There are two broad types of information that can be collected, called **quantitative** and **qualitative** data. Each has its own advantages and disadvantages, and gathering the right combination of the two is what leads to an overall successful scouting system.

Quantitative Data

Quantitative data is defined by numbers and figures. Examples of quantitative data include the number of basketballs shot into a goal, the number of frisbees fed in from a station, the amount of time it takes to chase down a scoring object, such as a ball, or the height of a stack of totes.

Quantitative data is the single most vital aspect of data collection. Scouting systems should place their primary focus on collecting quantitative data, because it is the easiest to interpret and gain real, impactful knowledge from. You can use quantitative data to calculate averages, max scores, minimum performance, differences in performance from match to match, trends, etc. all with valuable numbers, as compared to the nominative differences given by qualitative data. Quantitative data will always be the first source of information about a team's performance. The physical numbers do not lie.

However, quantitative data also has its drawbacks. It is difficult for human's to process quantitative data without the aid of a computer. It is even more difficult still to quantify certain aspects of robot interaction, such as how friendly their drive team is to how well they follow orders, or even how well a robot plays defense. We would like to be able to quantify these

things to capitalize on the advantages of quantitative data, but sometimes it is impossible. Sometimes, we can approximate a quantification through data processing, but this is not always the case.

Qualitative Data

Qualitative data is defined by words, images, and descriptions. Examples of qualitative data include how comfortable a drive team feels balancing a bridge with another team, how well a team can execute dodge maneuvers through a pyramid, whether a drive team listens to a game plan from cycle to cycle, or how comfortable a team feels scoring on a step.

All of these examples can technically be quantified, but the information is not the same. For example, someone looking for a good bridge balancer could look at the time it takes that robot to balance on average. Or dodging could be counted and given a score based on the number of times they dodged to the number of times they were blocked. Listening to a game plan could be reflected by the number of cycles completed in a match, and comfort scoring on a step could be reflected by the number of times the team had done it in the past. But these aren't exact quantifications, and they only approximate the qualitative information that accompanies them. Things like - "I felt confident balancing with this team" or "This teams coach yelled at me while balancing, which was too stressful" are helpful when deciding whether a team is good to play with, or what to do and what to avoid with that team.

Qualitative data takes precedent when information cannot be quantified properly. However, it has many drawback. Qualitative data is very subjective. If detail is left out, the same statement can mean many different things to many different readers. It is left to the person reading the data to interpret it as they read, instead of having it universally interpreted and giving everyone the same results. The previous statement - "coach yells, stressful" could be legitimate - the coach yells and its very stressful for that driver. But another driver may have no issue with someone yelling at them while they drive the robot. In that case, this data doesn't apply, but it wards people off from the team immediately.

Finding the Right Combination of the Two

It is important to quantify whatever information possible, and then any other information which is absolutely necessary to know about a team be collected qualitatively. Certain collection methods are inherently a combination of the two, such as Performance Maps.

In 2014, FIRST Team 701: The RoboVikes introduced a Performance Map they called the HotSpot system. This system would track where a robot shot a ball into a goal, over the truss, made or received assists, or caught a ball in relation to the field. This method both quantified data - the number of each of these occurring - and qualified it - where those actions were occurring. This provided a method of interpretation like no other before it for the team. One quick look at a team's HotSpot map, and you knew exactly where they liked to shoot over the truss from, exactly where they shot into the goals most often, and you knew where to go to play defense on them. While it does take some human interpretation, the dots are physical and hard to misinterpret. You aren't going to see a map with 50 successful shot dots clustered in the same spot and think - "I think they are bad at shooting from there." In this way, qualitative data can be combined with quantitative data to create a powerful method of interpretation.

First Steps - What Information is Important

The first step to collecting successful data is knowing exactly what information is important to you. Then, you turn those ideas and concepts into one of the two types of data, preferring quantitative. For example.

"The game involves shooting basketballs and balancing on bridges. Let's figure out what we want to know about the teams that perform.

- how many points did they score
- how accurate were they
- how quickly can they balance the bridge
- are they good at defense
- where do they shoot from
- how often did they balance"

Now, this guide will go over each stage of scouting the game during the match and what to look for in detail. This is simply an outline to look at what information may be important when playing with or against. Information like this can be turned into quantitative data like this -

" - number of shots they successfully shot (later processed into a score)
- number of shots they attempted to shoot (used with the previous for accuracy)
- number of balls they picked up (alternate method to calculate accuracy)
- timer on the balance of a bridge
- qualitative description of defense, paired with a checkbox of if they played defense or not
- either - list of places to shoot from and a count of each time they shot from that location OR a map and HotSpot method (described earlier in this document)
- count of successful and / or attempted balances"

Now, what makes certain data more important than other data? That answer is entirely up to what is important to you as a team and what strategy you play. A base of understanding would be the following.

- Scout ANYTHING that earns a team in game points, this is used to calculate an offensive score for the team.
- Scout MOST THINGS that are involved in maneuvering the field. For example, locations of scoring positions, ability to maneuver around or through any in game obstacles, ability to pick up objects from different locations on the field (feeder vs floor, etc.)
- Scout NOTHING that is extraneous to the data you need to know about yourself and other teams. Including too much information confuses scouters and makes it impossible to keep track of everything (such as, in 2015: number of cans knocked over during tele, method of can grabbing off the step, number of totes picked up off the step (this never happened anyway): If you Scouted for Team 701 in 2015, you know what its like to be looking at too much data).

Collection Methods

While many methods have evolved to collect data, they generally fall into two major categories, **paper scouting,** and **electronic scouting,** or some combination of the two.

Paper scouting is exactly what is sounds like. Students (or parents) sit in the stands, watching matches and recording information about robots on paper sheets. At the end of each match, a Scouting Director (or group of people) collect the paper sheets and compile the data by hand, either into a notebook, a master sheet, or some electronic method. Data collected by hand through Paper Scouting has its advantages and disadvantages.

With paper scouting, data collection is easy to implement. Simply arrange a sheet to collect the information you need, and print out as many copies as you need. Paper scouting is also relatively cheap in the short term, making it ideal for teams who cannot afford a large upfront cost for electronic devices to scout on. Paper scouting is also massively inefficient. It uses large amounts of paper and requires very good organization of documents in order to run well.

Electronic scouting uses some sort of computer system to collect data. Students possibly sit in the stands with tablets, laptops, or even game controllers and collect information about teams as they perform in matches. The data is processed by the computers through hard wire or wifi (sadly, WiFi is very difficult to maintain at competitions due to FRC rules), and interpreted into an excel document or some other database methodology. Data collected by computers is the preferred method of Team 701, but the advantages and disadvantages may pertain differently to each team.

Electronic scouting has its advantage in quick and easy data processing and interpretation. Collection is also more natural and sustainable, because it does not require a large amount of paper to handle. However, electronic scouting is difficult to implement, requires someone (or a group of people preferably) skilled in programming of some sort other than just the robot, and costs quite a bit of money to start-up (due to the cost of electronics). The nice thing about those costs is that they are a one time fee that can be covered through fundraisers, grants, etc. and do not recur at each competition each year.

Whether you choose to proceed with paper or electronic collection methods, the fundamentals of this guide still apply to you. Continue reading for best practices of processing and interpretation.

A Word on Pit Scouting

Collecting data through the pits is a good way of gathering preliminary information about teams. However, once you have data backed up by match performance, it is important to rely on pit scouting less and less. Trust not what teams assure you they can do in the pits, but what their robot's performance assures you they can do through real performance. As always, be courteous and maintain gracious professionalism. I won't go much further into pit scouting because I have little experience with it, but some teams swear by it. Personally, I much prefer the data that has been proven by real performance numbers than the claims of a team that wants to sound good on paper. While most teams are honest in my pit scouting experience, there is always a chance. Pit scouting is also a good time to go around at meet the teams and take pictures of their robots, which can be useful impressions for later on in the competition. The same principles apply to pit scouting when deciding what info to collect. Just get what is important to you. In my rookie year, we built a robot that would stack another robot on top of us, so we needed to know the size of the wheelbase and center of mass of other robots to determine their candidacy to balance on us. Just collect what seems important to you and your team.

Organizing Data Collection

In any collection method, it is important to use a system of ordering the data to make collection and processing easier and more intuitive. The easiest way I've found to organize data is through the inherent order of the match, which usually flows like this:

Pre-Match Information (stuff like: what team are you scouting? what alliance are they on? what does their robot look like? did they even show up?)

Autonomous Information (stuff like did they complete the autonomous tasks, did they move, did they have an autonomous strategy, etc.)

Tele-Operated Information (scoring, defense, etc.)

End-Game Information (for the last two years, has been included with the tele-op, because there was no end game. However, if an end game challenge is presented, this is where you record what the robot did to address it)

Post-Game Information (things like fouls, red/yellow cards, assistive play, defensive play, and any extraneous data that doesn't fit elsewhere and is easy to record at the end of the match)

When setting up data collection, keep these categories in mind. They are a good way of organizing electronic data or even paper sheets, and can help make processing easier (for example, have one person hand process the sheets for pre-game and auto info, then pass to the next for tele, to the next for end-game and post-game, etc.

Understanding Data Processing

Data processing is the simplest step of the system, but can also be the most important. Being able to process data quickly into the interpretation stage is vital to success as a scouting system. The most preferred method of processing is via computers, which are faultless in their processing powers and much, much faster than processing by hand. This stage is why collection using electronic tools is so advantageous. However, it is still possible to process hand collected data using an electronic method (ask around the members of Team 118: The Robonauts. Some of my good friends there showed me their fantastic picture processing design). Hand processing is slow and tedious, but it is still possible. One distinct advantage of hand processing is that it is much easier to process data match by match, thus staying up to date with the current matches that are playing. While this can be implemented through some work on the electronic end, many electronic systems run into trouble with WiFi due to recent (2011) FRC rules which ban such networks from event sites.

The best methods of processing data will be able to keep the interpreted data as up to date with real time as possible, which often means updating after each match, or after each 2 or 3 or some number of matches. Do what is possible for you, and try to collect data as often as it would warrant useful. If you don't have another qualification match until 4 hours from now, making decisions based on 4 hour old data isn't going to be helpful. But trying to make a decision in a matter of less than a minute based on the last match's data is also not helpful. Usually collecting data about 7-10 matches before your match gives you time to speak with the other drive teams and come up with a strategy based on recent data, if you cannot collect data after every match.

Once data is processed from the collection devices, whatever those may be, it needs to be transferred from the stands to the pit. For best use with a paper system, keep two copies of each team's master sheet, updated as time permits. Have someone run the needed sheets down to the pit and update the master sheets there and then take their copy back to the stands for later use. Because of the way match scheduling works, you will have at least 3 matches before you need to be back to the stands with the teams master sheets. With an electronic method of interpretation, this becomes much easier. Have a copy of your electronic device in the pits and the stands. Run the data down to the pits when your drive team needs it, or about every 10 matches. Your scouting director is a good resource in the stands for the drive team to call and ask about help with preparing the next match based on their thoughts and their own look at the copy of the data they have in the stands with them. This is probably the Scouting Directors most important job after keeping scouters in line and making sure they are collecting accurate data.

Understanding Data Interpretation

Data interpretation is the process of taking processed data and turning it into something that humans can easily read and understand. This includes things as simple as churning out comments for qualitative data to calculating averages and scores. Data Interpretation is as complicated as you need it to be, and is most popularly done through Microsoft Excel or a similar program. The program has a lot of processing capability, and can be very effective. But at FIRST Team 701: The RoboVikes, we use a very different method of data interpretation, something we affectionately call the **Chadabase** (after it's original programmer Matthew "Chad" Eatman).

The Chadabase

The Chadabase (yes its a pun on database) is the customized statistical analysis database used by FIRST Team 701: the RoboVikes to interpret data. It is fully customizable by programming through Python and pyTK tools, and gives a wide plethora of data analysis. FIRST Team 701 plans on releasing a public version of the Chadabase for teams to use in conjunction with their own scouting systems, but this is a project that has been in the works for nearly 3 years and isn't anywhere near completion. As an example of interpretation methods, the Chadabase offers the following.

Team Information: The Chadabase is fully equipped to give every detail imaginable about a team's performance at an event. From max, min, and average scores to linear trends of performance and graphical displays of almost every piece of data, the Chadabase offers unparalleled detail and interpretation of team information on a micro-scale.

Rankings: The Chadabase uses algorithms designed to rank teams on a multitude of criteria, from scoring of individual objects to overall scores, to defensive play to fouls. It also includes a system for weighting scores based on the difficulty of the opposing alliance (when there is one) and places a heavier weight on scores that occur towards the end of qualifications (as they are a better reflection of how the team will perform in eliminations.

Search Options: The Chadabase can search through every team at an event for nearly any possible criterion that you could imagine, from autonomous score to number of scoring objects picked up from the ground to variations in a fraction of a second of pickup times. This method is best used to create a pick list based on your criteria. You can save and sort your pick list within the Chadabase itself, allowing for easy access during alliance selection.

Match Predictions: Not to be left out, the Chadabase incorporates statistical confidence methods to predict the outcome of real or hypothetical matches, and then gives you the information about why the match with turn out the way it does. Not as an end-all, do-all, the system is meant to show you whether you should play a match to your strengths or capitalize on your opponents' weaknesses to pull out a victory. It is also useful to compare total expected scores for different plausible elimination alliances to help you with those pick lists.

Alliance Selection Algorithms: You can enter and save the elimination alliances straight from the Chadabase itself, following alliance selection as it happens and recalling the information later for Match Predictions. Most notably, the Alliance Selection screen keeps a dynamic pick list, graying out teams who have already been selected and helping you keep track of which teams are left.