Reliability Analysis of Mobile Robots

Jennifer Carlson, Robin R. Murphy
Center for Robot-Assisted Search and Rescue
Computer Science and Engineering
University of South Florida
Tampa Florida

***Abstract*— Failure data on mobile robots is critical for three reasons: to support the theory of autonomous fault detection, identification, and recovery necessary for success in new domains; provide design and manufacturing feedback to the robotics community; and permit project managers to accurately create development schedules. The failures considered in this paper occurred over a period of 2 years, in a variety of environments. Failure type and frequency data were collected from thirteen robots representing three manufacturers and seven models. The data was analyzed using standard manufacturing measures for the reliability of a product. The *mean time between failures* represents the average time to the next failure. *Availability* was used to gauge the impact of a failure on a project. The results show that the reliability for mobile robots is low, with an average MTBF of 8 hours and availability of less than 50%. The platform itself was the source of most failures (42%) for field robots, while the control system was responsible for 29% of the failures.**

I. Introduction

Mobile robots have been an active topic of research since 1967.[9] They are becoming increasingly important for civilian applications such as urban search and rescue (USAR), and military applications such as military operations in urban terrain (MOUT) and the Future Combat System (FCS). Failure data on mobile robots is critical for at least three reasons: to support the theory of autonomous fault detection, identification, and recovery (FDIR) necessary for successful fielding in these challenging domains; to provide design and manufacturing feedback to the robotics community; and to permit project managers to accurately create development schedules.

In order to autonomously diagnose and recover from robot failures, causal models such as [7] and [18] require the type and frequency of failures. The failure mode versus the symptom is also important in designing cooperative diagnostic systems such as SFX and teleVIA. The frequency of failures of components under what conditions are important information for manufacturers. They can use this data to make improvements to the current generation of robots and to be proactive in their designs of the next generation. Failure data is needed to allow researchers and program managers to better estimate, and control, development time. For example, if robots are unavailable due to repairs for 25% of the year, then projected schedules must take that into account in order to be accurate or a technician might be hired to complete repairs faster than an average software specialist would be able to do so.

The experience at the University of South Florida (USF) with mobile robots provides a reasonable preliminary database. The Center for Robot-Assisted Search and Rescue, formerly the Perceptual Robotics Laboratory, has nineteen mobile robots from five manufacturers. These robots are either research robots or commercially-available robots built in limited quantities. Research at USF tests software under laboratory conditions, and/or high fidelity outdoor simulations of military operations in urban terrains (MOUT) and urban search and rescue (USAR). More than 100 hours of field work per year are conducted by the lab. Ongoing research has been conducted in the area of sensor failure detection, identification, and recovery since 1993 [4], so there is significant motivation to observe and analyze failure data.

This paper examines the user logs of the most heavily used robots at USF and collected failure type and frequency data for each subsystem of a mobile robot: *effector* (or platform), *control system*, *power*, *wireless communication*, and *sensing*. In addition, the failure data was further analyzed using standard manufacturing measures for the reliability of a product. The *mean time between failures* represents the average time to the next failure. *Availability* was used to catalog the impact of a failure on a project.

II. Review of the Literature

Failures are not an uncommon topic in the literature though most work in robotics focuses on fault detection and diagnosis. In this area, model-based methods are predominant. Such methods use explicit models of the normal behavior of the robot to detect and diagnose fault conditions. Examples of these can be found in [6], [7], [17], and [18]. Other methods are based on rules at the planning level or partial casual models of the system. Representative examples are [2], [3], [4], and [15]. Less common methods for dealing with failures on robots are expert systems [5] and agreement-based systems [14] which use only agreement between redundant sensors to detect faults.

What is not found in the literature is analyses which explore how robots fail and the underlying reasons behind those failures. One reason for this may be that the design of a fault tolerant system for mobile robots is focused on the ability to handle or at least to learn from new, un-modelled faults. While this is feature is important, knowledge of common failures, their causes, and repair strategies can help in the design and optimization of such systems. The designers of systems found in [2], [3], [5], the planner for [15], and [18] would be far more successful with accurate information about how robots fail. Mackey's [7] system is already designed to use whatever failure information its developers can provide while only slight modifications to [4] would allow it to use frequency information to optimize the diagnosis process.

A recent workshop on Robots in Exhibitions produced two papers which provide some insight on the reliability of mobile robots actively used for long periods of time. Nourbakhsh [12] describes a set of four autonomous robots used for a period of five years as full-time museum docents. Their robots reached a mean time between failures (MTBF) of 72 to 216 hours. Tomatis [16] is especially interesting in that their results are similar. While their analysis is more narrow in both the applications and robots analyzed, their MTBF was 7 hours, similar to the 8.3 MTBF found here. Also, they did not break down their failure analysis into the same sets of components, but it is interesting to note that the component mentioned as being particularly troublesome was a motor controller which is consistent with the findings described in Sec. IV-B.

Another instance of this type of analysis in the robotics literature is [8], which is a detailed analysis of the failures encountered with the robots used during the World Trade Center (WTC) rescue operation. This paper expands the knowledge of how mobile robots fail by analyzing the data for failures which occurred in a variety of environments and over a long period of time (2 years), as compared to the two week period during the robot-assisted WTC response [8].

III. METHOD

User and failures logs served as the sources of data for this analysis. A total of 97 failures were recorded over a period of two years, specifically June 21, 2000 through July 31, 2002. Prior to February 2002, informal records were kept including changes to the robots and information about ongoing repairs. Starting in February 2002 formal failure and user logs were kept. The user logs were entered by the robot's operator and the failure logs were recorded by the person who performed the repair. Since then over 670 hours of usage have been logged, including 283 hours of field work.

The following information was gathered for quantitative analysis: *which robot was involved*, *who repaired it*, *the date the failure was discovered*, *the date the failure was fixed*, *the total repair time*, *which component failed*, *where the failure occurred*, and *where the repair was performed*. In order to determine the expertise required to effect a repair, the individuals who performed the repair were separated into two groups: hardware specialists and operators. A similar taxonomy was created for the location of the failure and the repair; see Sec. III-B.

It should be noted that there is a gap of 5 months in which no failures were recorded. Much of this time was spent writing up the results of earlier field work, and participating in the rescue and recovery efforts at the World Trade Center. The majority of the robots used at the WTC were sent to the manufacturer for repair, and were not returned for several months. Therefore the majority of robots used in the lab were not available during that time period. As mentioned in Sec. II, a detailed analysis has already been done on the use of robots at the WTC response, using video footage.[8] The failures identified during the response are not included in this analysis.

A. Robots

Of the nineteen mobile robots at USF, thirteen were considered by this analysis. These thirteen robots represent seven different models made by three manufacturers. They are used for funded research as opposed to the remaining six which are dedicated to education and demonstrations. Robots mass produced for the entertainment industry, such as the Sony Aibo, are not included. Eleven of the robots serve in field domains. Field robots are expected to work outdoors, though generally not in rain or snow. They are intended to be able to handle rougher terrains, tolerate dirt and dust, even multi-story falls. The two indoor robots are the more traditional research robots, with small, narrow wheels suitable for operating on smooth flat surfaces.

The intention of the paper is to provide a useful quantification of how robots fail, not to compare and contrast the reliability of one manufacturer or model to another. Indeed, the results below suggest that field robots have similar reliability scores regardless of manufacturer. To maintain focus on how and how often robots fail rather than which robots fail, the paper labels the three manufacturers by *X, Y,* and *Z*, with models labeled *A*... *G*.

Table I includes the label for the robot's manufacturer and model as well as the number of robots used in the lab, the robot's size, and the general application for which it was designed. The size of a robots is either *man-packable* or *man-portable*.[8] A man-packable robot can be safely carried by one person. A man-portable robot is larger than a man-packable robot but still can be carried in a HUMMV or personal car and can be lifted in and out by one or more people. A Remotec Andros robot for Explosive Ordinance Disposal is an example of a mobile robot that is *not* man-portable; it requires a special trailer and lifts or ramps to transport it.

A robot is usually designed for a specific domain. Models A and B were designed for chemical and nuclear inspection, though they were used for urban search and rescue (USAR) and military operations in urban terrains (MOUT). Models C and D were specifically designed for MOUT, while E and F were designed for general outdoor research. Model G was intended for indoor research.

Field X A and B model robots are shoebox sized robots with a footprint no larger then 15.5 by 30.5 cm, see Fig. 1. Both are tracked vehicles and do not have onboard computers. Both have the same onboard sensor suite, which consists of a microphone, speaker, a motor-driven manual-focus CCD camera, and a camera tilt unit with halogen lighting. Model B robots also have the ability to adjust the shape of the platform's chassis to raise or lower the camera tilt unit and change the track profile.

TABLE I

THE ROBOTS USED AND SOME OF THEIR CHARACTERISTICS.

Model	Size	Manu.	#	Purpose
A	man-packable	Field X	1	inspection
B	man-packable	Field X	3	inspection
C	man-packable	Field Y	3	MOUT
D	man-packable	Field Y	2	MOUT
E	man-portable	Field Y	1	outdoor research
F	man-portable	Field Y	1	outdoor research
G	man-portable	Indoor Z	2	indoor research

Fig. 1. A Field X man-packable inspection robot.

Field Y C was a precursor of the D model robots. Both are about the size of a large backpack with a footprint of 61 by 51 cm, see Fig. 2. They are tracked vehicles with mobile Pentium II or III class processors onboard. Both the C and D models carry multiple cameras and lighting. The Model C robots also have a set of 13 sonar range sensors. These robots can also be modified to carry two-way audio, thermal imaging, low-light cameras, GPS, attitude sensor, compass, chemical detectors, and manipulators. Both were developed for MOUT operations though only the Model D has features which make it durable enough for such operations. For example it is shock resistant, water resistant, and less susceptible to detracking.

Field Y E and F models are larger, wheeled robots with skid steering. The E model robot has a footprint of 78 by 62 cm and the F model is 104 by 81 cm. Both carry mobile Pentium II or III class processors. Both the E and F models carry multiple cameras. These robots can also be modified to carry two-way audio, thermal imaging, low-light cameras, GPS, attitude, compass, chemical sensors, and manipulators. The E model robots are small enough to be used for both indoor or outdoor research projects. The larger, Model F, robots are less maneuverable, but have a much longer battery life and the capability of carrying smaller robots like the Model C or D's.

G model robots, shown in Fig. 3, are cylindrical in shape, with a 53 cm diameter. Both are wheeled robots with synchronous, non-holonomic drive systems. These robots have two onboard computers with Pentium-class processors. Their sensor suite may include tactile, ultrasonic, and basic vision systems.

Another important factor to consider when comparing robot models is their maturity. The Field X robots are the most mature; over ten years worth of experience with

Fig. 2. A man-portable general purpose field robot (top), and a MOUT field robot interacting with a larger field robot (bottom), all from manufacturer Field Y.

Fig. 3. Two models of Indoor Z's research robots.

similar platforms went into the design of these robots. The G model was developed in 1996. Both E and F models have been in production for about five years. The C model robots were first developed in 1999 and went through several major modifications during the next two years. The D model is the newest, it was first produced in 2001.

B. Definitions

For the purposes of this paper, a failure is defined as *the inability of the robot or the equipment used with the robot to function normally*. Both complete breakdowns and noticeable degradations are included in this analysis.

An example of a failure that was encountered was a problem with the wheels on the Model C robots. In two cases, the wheels warped due to exposure to heat in field exercises, immobilizing the robot. An example of a common failure for Field X robots is a failure of the control system where the robot becomes unresponsive, or freezes. This may be an example of support equipment failure if the source of the problem is not on the robot, but within the control unit used to tele-operate the robot. A good example of a degradation was encountered with a Model G robot. In this case, a faulty camera cable caused signal loss from a camera on the robot. The rest of the robot, including the second camera, was not affected by this failure. Such degradations may or may not affect the usability of the robot, depending on the task. A task which requires stereo vision, for example, could not be performed with a single camera.

Each failed component was separated into one of five categories, these being *control system, effector, power, sensing*, and *wireless communications*. The control system category includes the onboard computer and software or the control unit. Effectors are any devices that performs actuation such as the motor, appendages, treads/wheels, and any connections related to those components. The power category includes any component that affects the power system of the robot. Examples of these would be the batteries, chargers, and various connections allowing the robot to be powered. The sensing category includes any sensors and the connections or software that may affect their functioning. Finally, the wireless communications category covers the wireless equipment used by Field Y and Indoor Z.

The location in which the failure occurred was placed in one of two categories: *in-lab* or *in-field*. In-lab failures are simply any failures which occurred in the lab. In-field failures occurred outside of the lab, usually during demos, outdoor testing, or training sessions.

The type of failure is similarly divided into a binary category of *field-repaired* versus *not field-repaired*. Field-repaired failures were failures which were repaired in the field. These repairs may have or may not have required a hardware specialist to carry out the repair. Note that field repaired is not the same as field repairable. All other failures fall into the not field-repaired category. These include failures which were fixed in the lab and those which were returned to the manufacturer for repair.

C. Calculations

All the formulas used for reliability analysis of the data were taken from the IEEE standards presented in [13]. The mean time between failures or MTBF is calculated by equation (1). This metric provides a rough estimate of how long one can expect to use a robot without encountering failures. Another metric used in this analysis is the failure rate, which is simply the inverse of MTBF. Availability is calculated using (3), where the Mean Time To Repair, *MTTR* is defined as in (2).

$$MTBF = \frac{\text{Number of Hours Robot Was in Use}}{\text{Number of Failures}} \quad (1)$$

TABLE II

OVERALL FREQUENCY AND MTBF BROKEN DOWN BY MANUFACTURER.

Manufacturer	Number of Failures In-Lab	Number of Failures In-Field	Failures/hr	MTBF (hrs)
Field X	3	34	0.17	6.03
Field Y	31	12	0.16	6.13
Indoor Z	16	0	0.05	19.50
Overall	50	46	0.12	8.29

$$MTTR = \frac{\text{Number of Hours Spent Repairing}}{\text{Number of Repairs}} \quad (2)$$

$$Availability = \frac{\text{MTBF}}{\text{MTBF + MTTR}} \cdot 100\% \quad (3)$$

It may be recalled that the usage logs do not cover the entire time-frame in which the failures occurred. In an attempt to remedy this discrepancy, logs were added for every failure which did not already have a corresponding entry in the usage logs. The estimated usage hours for the added logs was determined by taking the average duration of similar, known entries. Similar usage entries correspond to logged failures for robots of the same type. If no such records exist, all the entries for that robot type were considered.

Any other values included in this analysis were calculated using the implied formula. For example, the probability that a failure was caused by a component type c is simply (4).

$$P(c|failure) = \frac{\text{Number of Failures Caused by c}}{\text{Total Number of Failures}} \quad (4)$$

IV. RESULTS

This section examines the frequency of failures, the probability that a failure was caused by a particular type of component, and the impact of failures in terms of availability and downtime.

A. How often do robots fail?

Table II shows how frequently failures occur with the robots in Table I. It shows the total number of failures recorded divided into in-lab and in-field as defined in Sec. III-B. It also shows the overall frequency of failures, in failures per hour, and the mean time between failures (MTBF), in hours. The failures are grouped by manufacturer, with overall statistics provided at the bottom. One of the Field Y failures actually occurred while the robot was with the manufacturer. This location could not be categorized as either in-lab or in-field, leaving the total number of failures at 96, as opposed to 97. This failure was included in the overall frequency and MTBF calculations for that manufacturer.

This table provides a high level prospective on how often robots fail and under what conditions. It also raises some additional questions. The answer to those questions lays in the more subtle factors hidden in the underlying data and in the characteristics of the robots themselves.

Oddly, the Field Y robots have had more failures in the lab then in the field. This is apparently because they are used more in the lab, only 30% of the logged usage time is in the field. This statistic is largely influenced by the Model C's which rarely leave the lab because they suffer from severe failures which make them impractical for use in the field. Examples of the failures seen, while the robot is still in the lab, include the battery shorting, the mercury switch shorting, and failure of the power hub which leaves the robot immobilized.

Field X robots have had more failures in the field then the Field Y, does that mean the Field X robots are less reliable in the field? Field X robots tend to be used more often in the field, 94% of the logged usage time is in the field. The reason for this is that the failures these robots suffer from are less severe and easy to repair in the field. The two most common failures are the tracks coming off and the control system freezing. Both take less then five minutes to fix and do not require a hardware specialist's expertise. Based on these considerations, Field X robots have had far more opportunities to fail while in the field, both because they are used more often and because the failures are short-lived. This is the most likely reason for Field X robots to have more failures in the field then Field Y's. The fact that their MTBF rates are similar would suggest that they are equally reliable.

Why do the Field Y and the Field X robots fail more than the Indoor Z research robots? The failures the Model G's have experienced over the past couple of years have either been failures of supporting equipment, like the robot's charger, power supply, or wireless Ethernet base units. The fact that they are only operated in the indoor environments for which they are designed is a large contributing factor. It is also important to keep in mind that most of the field robots have innovative capabilities like self-righting and shape-shifting. The frequency of failures for these new models, especially as compared to the research robots, would suggest that more development and testing is needed before these robots can be trusted to operate in their target environments.

B. *Which components fail?*

Table III was generated using the component categories defined in Sec. III-B. As in the previous table the failures are grouped by manufacturer with the overall probabilities for each category shown at the bottom of the table. The wireless communications category does not apply to Field X robots because all of them are tethered to the control unit.

Effector failures tend to be the most common. Over a third of these failures are treads coming off of their tracks on the Model B's or Model D's. Other examples of effector failures would be the wheels warping on the Model C's, Model B's pinion gear becoming stripped, and the failure of a motor-amp on the Model E. The type of drive appears to have the largest effect on the number of failures in this category. Effector failures are the most common for all but Model A of the tracked robots. Only 3 effector failures occurred on wheeled robots.

TABLE III

PROBABILITY THAT A FAILURE WAS CAUSED BY A COMPONENT TYPE BROKEN DOWN BY MANUFACTURER.

Manufacturer	Effector	Control System	Power	Comms	Sensing
Field X	0.49	0.35	0.05	N/A	0.11
Field Y	0.36	0.30	0.16	0.09	0.09
Indoor Z	0	0.13	0.25	0.50	0.13
Overall	0.35	0.29	0.13	0.12	0.10

The second most common source of failures is the control system. In over half of these cases the robot was unresponsive and the solution was to cycle the power; the source of these problems remains unknown. Other examples of control system failure include a corrupted hard drive on an Model C, a timing delay which hung the boot process on the same Model C, and a lag in a Model B's response to left turn commands. Again there is a clear difference between the field robots' and the Model G's control systems. Again this is probably due to the fact that the field robots are operated in far more challenging, unstructured environments.

As might be expected, about half of the power failures on the robots are due to the battery or its connections to the robot. Other examples of these types of failures include a power switch on a Model C which did not make contact and a wire contact broken inside a Model G's charger. The fact that Field X robots' batteries are carried by the operator instead of the robot (they are connected to the robot through the control unit and tether) might explain why there are so few power failures for these robots as compared to the others.

The wireless communications category is a tight group including mainly problems with Field Y's or Indoor Z's wireless equipment. The predominant failure is unexplained communication loss. All of these failures occurred within the lab. The fact that the Model G's experience more of these failures then the other robots is probably due to the fact that it uses Radio-Ethernet, which was a precursor to all the modern wireless standards.

According to the data, the least common source of failure for these robots is sensing. The source of half of these failures was broken connections between the sensors and the control system. The most common failed sensor is the camera. Other examples of this type of failure include a faulty sonar cable and a damaged compass, both found on Model C's. These failures appear to be similar and equally uncommon among the robots used in the lab. This is due in part to the fact that the manufacturers purchase mass produced sensors to install on the robots. Conversely, the

TABLE IV

COMPARISON OF THE PERFORMANCE OF RESEARCH VERSUS FIELD ROBOTS. ONLY FAILURES IN THE TARGET ENVIRONMENT ARE INCLUDED.

Manufacturer	Type	# of Failures	% of Usage	Failures/hr	MTBF (hrs)
Field X	Field	34	94%	0.16	6.14
Field Y	Field	12	28%	0.16	6.27
Indoor Z	Research	16	100%	0.05	19.50

TABLE V

AVERAGE DOWN TIME AND AVAILABILITY.

Manu.	Availability	Average Downtime (hrs)		
		Overall	Field Repaired	Not Field Repaired
Field X	83.6%	195	0.13	703.3
Field Y	23.9%	353	2.66	411.3
Indoor Z	93.9%	60.5	0	60.5
Overall	46.6%	243	0.6	367.9

robot's effectors, control, and power systems are custom built by hand.

C. *Performance of research versus field robots*

In order to compare research and field robots it is important to consider only failures which occurred in the environment for which each robot was designed. The statistics in Table IV are calculated using only failures which occurred in the robot's target environment. For research robots that is only in-lab failures, and for field robots only failures which occurred outside of the lab, are included. For reference, the number of in-lab or in-field failures is copied from Table II. The percentage of usage in the target environment over all the recorded usage is also included. The performance in terms of failure metrics is captured in the overall frequency of failures and the mean time between failures (MTBF).

The statistics presented in this table indicate that the research robots fail less often in their target environments than field robots. The MTBF for the research robots is more than three times the MTBF for both Field X and Field Y's field robots. The number of failures provides a less clear picture. Field Y's models have failed fewer times in their target environment then Indoor Z's. It would be misguided to consider this a point in the field robots' favor, considering that these robots fail more in the lab (72% of failures are in lab) than in the field. The failures and usage data recorded to date lead to the conclusion that the research robots are more reliable in their target environment then the field robots. As mentioned in Sec. IV-A, this appears to be due to the innovative capabilities of these robots, and the inherent difficulty in constructing a robot which can operate in unstructured, outdoor environments. Again the data would suggest that more development and testing is needed for these robots to attain the level of reliability seen with Indoor Z's research robots.

D. *Impact of failures*

Table V shows the collective impact of these failures. The projected availability of the robot is included as a percentage of time. This metric, also called reliability, should be interpreted as the probability that the robot will be functional at a particular point in time. The average downtime, divided into overall, field-repaired, and not field-repaired as defined in Sec. III-B, is also included. Average downtime is the average amount of time in which the robot was not used as the result of a failure. The failures are again grouped by manufacturer and then summarized at the bottom. The average downtime for field-repaired failures on Model G's is 0 because no failures of this type have been recorded.

In studying these numbers it is important to remember that availability (3) is calculated using the average repair time (2). Other factors which affect the average downtime, like low priority repairs waiting while other tasks are completed by the hardware specialists, can make it much larger then the average repair time. This influence is most evident in the availability of the Field X's, which is quite high considering that they have the highest average downtime for failures that were not repaired in the field. It also helps that 70% of the failures recorded for the Field X's were repaired in the field, and these failures have both low repair and downtimes. The opposite influence can be seen with Field Y's models where 86% of the repairs are carried out in the lab or require returning the robot to the manufacturer. The repair time for these failures is customarily longer which results in the low availability rate seen. Another good example of the difference between downtime and repair time is Model G's average downtime of 60.5 hours, which is much larger than the 1.26 hours calculated for average repair time. In this case, two failures were left unfixed for over a week while the hardware specialists were on vacation, whereas the repairs for these failures only took a few hours each.

V. DISCUSSION

This section offers some support for the validity of the statistics presented in Sec. IV as well as a discussion of the implications of these results. Finally, some improvements for future data collection are presented.

A. *Validity of the data*

It may seem counter-intuitive, with an overall mean time between failures of less than 9 hours, that only 97 failures were recorded over a period of two years. Keep in mind that MTBF is often used for systems like power plants, and fileservers which are in constant use. This is not the case with mobile robots. Table VI considers only the time period for the user log, broken down by month, during which 673 hours (28 days) worth of usage is logged. The 'Field Usage' and 'Total Usage' columns show the

TABLE VI

USAGE AND FAILURES BY MONTH FOR 2002.

Month	Field Usage	Total Usage	# of Failures
Feb	24.2%	38.1%	7
March	4.6%	11.8%	12
April	0.6%	10.9%	4
May	17.1%	33.4%	2
June	2.9%	5.2%	1
July	7.5%	18.7%	11
Overall	7.4%	17.1%	37

percentage of time that the robots were used in the field and overall, respectively. The last column shows the total number of failures recorded. The last row provides a summary of the same information over the entire 5 month period. This shows that the robots were only actively used about 17% of the time.

Even taking into account the fact that robots are not in constant use, over a period of 673 hours a total of 81 failures should have occurred whereas only 37 were logged. The reason for this disparity is that the 8.3 hour MTBF statistic was calculated from all of the failures, including those which occurred before usage was logged. Therefore the usage time associated with these failures is estimated. This discrepancy is offset, however, by the fact that common failures tend to go unrecorded. Good examples would be system freezes which can be fixed by rebooting. Some failures happen so frequently that experienced operators will often fix the problem without consciously realizing that a failure occurred.

Another point to consider is that over a third of the logged usage over the past two years has occurred in the field. In this case the field consists of a variety of environments, from demos for kids in carpeted classrooms, to naval research vessels, to urban search and rescue testbeds. The vast majority of field environments the robots have encountered were not scientifically managed for their safety, but qualify as real world experiments which test the limits of their capabilities.

B. Implications

The results indicate that field robots fail more often and are available less than research robots. One possible explanation is that the design and specification of robots for demanding field conditions is either too narrow or incomplete. The types of failures also imply manufacturing deficits, particularly with quality control.

Field robots are often used for domains outside of their original scope. For example, Field X robots are designed for inspection of pipes or ducts with hard, smooth surfaces. The most common failure for Field X robots is detracking while turning on coarse surfaces (e.g., carpeting, rubble). While it may seem unreasonable to use field robots outside of their specific domain, it is important to note that the Field X robots, which are used in practice in applications farthest from the original intent, have the highest availability (83.6%). Also, many field robots are expected to be multipurpose as the demands of the field vary considerably.

Design flaws appear to account for some of the failures. Some examples include wheels warping due to heat and batteries shorting due to lack of insulation. It should be noted that five models of the total seven considered by this study were prototypes undergoing rapid evolution; these models were expected to have design flaws. Consider that Model E from Field Y had the highest reliability with an availability rate of 95.8%, but Model C from the same manufacturer had one of the lowest availability rates with 17.1%. Model E is a larger vehicle based on mature technology and which has been in production for many years, while Model C was a radically new design.

There is also evidence that many failures stemmed from a lack of quality control during manufacturing. For example, robots have arrived with several wires pinched by the cover plates. However, establishing and maintaining quality control procedures is particularly demanding for small batches of customized products. A small company may not have access to specialists in manufacturing and quality control. This signifies the need for greater investment in the manufacturing process.

The US government pays much of the development costs of the field robots; it can also provide industry partnerships and manufacturing support. Another way for companies to get the needed capital for investing in quality control is to charge the consumers higher, more realistic, prices.

C. Improved data collection

The data analyzed in this paper was a useful preliminary database, but more data, and more types of data, is needed. Automating both the usage and the failure data collection would improve the accuracy and completeness of the data. The data which is desirable to record consistently in the future are: *the operational environment*, *the intended mission or task*, and *the symptom of the failure and the actual cause*. This additional information would help with both the analysis of under what conditions robots fail and how they are diagnosed.

VI. CONCLUSIONS

Based on 673 hours of actual usage by thirteen robots and three manufacturers, it appears that mobile robots, in a given hour, have a 5.5% probability of failure. The reliability is very low, with an average MTBF on the order of 8 hours and availability of less than 50%. As expected, field robots have higher failure rates and overall lower reliability than indoor robots, possibly because of the demands of the outdoor terrains and the relative newness of the platforms. The effectors, or platform itself, was the source of most failures for field robots whereas the biggest failure in indoor robots was with the wireless communication link. Surprisingly, the control system, either the hardware EEPROMs or the operating system itself, were responsible for 29% of the failures.

The reliability of a robot appears to be independent of manufacturer, and is most often related to the maturity and application of a specific model. As would be expected, new models fail more frequently than older models, as the design and manufacturing bugs are worked out, although the frequency and severity of the flaws seem high.

The failure data contributions to the FDIR and autonomic computing communities. Manufacturers should be alerted to the nature of the failures, particularly in the design of the platform and the installation of the control system. Also, manufacturers and users may need to plan ahead for repairs; either manufacturers need to provide more rapid turnarounds or users hire technicians to support on-site repairs. Users of mobile robots should allow for a less than 50% availability of the robots in their development schedules.

Current and future work is concentrating on creating even more detailed logs and on-board "black boxes" so that more quantitative and descriptive information can be gathered.

Acknowledgments

The work reported in this paper was supported by grants from the Department of Energy RIMCC grant, Office of Naval Research Grant N00773-99PI543, and the DARPA Tactical Mobile Robot award F04701-99-C-0317. The authors would like to especially thank Carson McCoy for gathering the basic data and Brian Minten for his helpful comments, and DIA for the donation of a robot.

VII. REFERENCES

[1] G.T. Chavez and Robin R. Murphy. Exception handling for sensor fusion. *SPIE Sensor Fusion VI*, pages 142–153, September 1993.

[2] B. Deuker, M. Perrier, and B. Amy. Fault-diagnosis of subsea robots using neuro-symbolic hybrid systems. In *IEEE Oceanic Engineering Society. OCEANS'98. Conference Proceedings*, pages 830–834, 1998.

[3] K. Djath, M. Dufaut, and D. Wolf. Mobile robot multisensor reconfiguration. In *Proceedings of the IEEE Intelligent Vehicles Symposium 2000*, pages 110–115, 2000.

[4] David L. Hershberger. Mobile robot sensor fault recovery improvements. Master's thesis, Colorado School of Mines, June 1997.

[5] M. Krishnamurthi and D.T. Phillips. An expert system framework for machine fault diagnosis. *Computers and Industrial Engineering*, 22(1):67–84, 1992.

[6] K.B. Lamine and F. Kabanza. History checking of temporal fuzzy logic formulas for monitoring behavior-based mobile robots. In *Proceedings 12th IEEE Internationals Conference on Tools with Artificial Intelligence*, pages 312–319, 2000.

[7] R. Mackey, M. James, Han Park, and M. Zak. Beam: Technology for autonomous self-analysis. In *2001 IEEE Aerospace Conference Proceedings*, pages 2989–3001, 2001.

[8] Mark Micire. Analysis of the robotic-assisted search and rescue response to the world trade center disaster. Master's thesis, University of South Florida, July 2002.

[9] Robin R. Murphy. *Introduction to AI Robotics*. The MIT Press, 2000.

[10] Robin R. Murphy and Ronald C. Arkin. Sfx: An architecture for action-oriented sensor fusion. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1079–1086, July 1992.

[11] Robin R. Murphy and Erika Rogers. Cooperative assistance for remote robot supervision. *Presence*, 5(2):224–240, Spring 1996 1996. special issue on Starkfest.

[12] Illah R. Nourbakhsh. The mobot museum robot installations. In *Proceedings of the IEEE/RSJ IROS 2002 Workshop on Robots in Exhibitions*, pages 14–19, 2002.

[13] Institute of Electrical and Inc. Electronics Engineers. Ieee recommended practice for the design of reliable industrial and commercial power systems. IEEE Std 493-1997, 1998. IEEE Gold Book.

[14] M. Soika. A sensor failure detection framework for autonomous mobile robots. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovative Robotics for Real-World Applications*, pages 1735–1740, 1997.

[15] E.R. Stuck. Detecting and diagnosing navigational mistakes. In *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, pages 41–46, 1995.

[16] N. Tomatis, G. Terrien, R. Piguet, D. Burnier, S. Bouabdallah, and R. Siegwart. Design and system integration for the expo.02 robot. In *Proceedings of the IEEE/RSJ IROS 2002 Workshop on Robots in Exhibitions*, pages 67–72, 2002.

[17] M.L. Visinsky, J.R. Cavallaro, and I.D. Walker. A dynamic fault tolerance framework for remote robots. *IEEE Transactions on Robotics and Automation*, 11(4):477–490, 1995.

[18] R. Washington. On-board real-time state and fault identification for rovers. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings*, pages 1175–1181, 2000.