

A Predictive Approach to Eliminating Errors in Software Code

Originating Technology/NASA Contribution

NASA's Metrics Data Program Data Repository is a database that stores problem, product, and metrics data. The primary goal of this data repository is to provide project data to the software community. In doing so, the Metrics Data Program collects artifacts from a large NASA dataset, generates metrics on the artifacts, and then generates reports that are made available to the public at no cost. The data that are made available to general users have been sanitized and authorized for publication through the Metrics Data Program Web site by officials representing the projects from which the data originated.

The data repository is operated by NASA's Independent Verification and Validation (IV&V) Facility, which is located in Fairmont, West Virginia—a high-tech hub for emerging innovation in the Mountain State. The IV&V Facility was founded in 1993, under the NASA Office of Safety and Mission Assurance, as a direct result of recommendations made by the National Research Council and the Report of the Presidential Commission on the Space Shuttle Challenger Accident. Today, under the direction of Goddard Space Flight Center, the IV&V Facility continues its mission to provide the highest achievable levels of safety and cost-effectiveness for mission-critical software.

By extending its data to public users, the facility has helped improve the safety, reliability, and quality of complex software systems throughout private industry and other government agencies. Integrated Software Metrics, Inc., is one of the organizations that has benefited from studying the metrics data. As a result, the company has evolved into a leading developer of innovative software-error prediction tools that help organizations deliver better software—on time and on budget.

Partnership

Since 2002, Integrated Software Metrics has not only studied NASA's metrics data, but has contributed to the maturation of the Agency's Metrics Data Program, through a contract with Galaxy Global Corporation, Inc.; both Integrated Software Metrics and Galaxy Global reside nearby the IV&V Facility in Fairmont. This contract enabled Integrated Software Metrics to work together with Glenn Research Center to generate metrics for the Metrics Data Program's data repository. NASA is now leveraging what was learned from this collaboration to better identify error-prone computer code and, hence, assure mission success.

Commercially, Integrated Software Metrics has tapped into everything it has learned from its partnership with NASA to create a new, artificially intelligent product suite

called Predictive. Prior to introducing the software to market, the company tested it on very large NASA software projects consisting of over a million lines of computer code, in order to ensure its efficacy.

Product Outcome

The Integrated Software Metrics Predictive suite of software products predicts where errors will occur in software code. Such a capability enables users to uncover any errors in the early stages of software development, thus saving time and money.

Software errors are usually not found until the late stages of the development cycle, when it becomes very costly to go back and fix them. Addressing these errors, however, is essential; otherwise, software developers build a reputation for delivering faulty products or, even worse, create life-critical situations when the software is part of larger systems or devices, such as power plant-monitoring systems or medical equipment.

Integrated Software Metrics is addressing these problems on the front end, before it becomes too late. The three products that make up its new Predictive suite are: Predictive Lite, Predictive Pro, and Predictive Server.

Predictive Lite was launched in February 2005 and provides basic metrics analysis and error prediction across one or more projects, for those who develop and manage software, including project managers, individual developers, and software quality assurance professionals. While identifying error-prone code, Predictive Lite will create color-coded reports that highlight risk areas. The user can then focus attention on the flawed code and assign resources to check and potentially fix the code before the project is completed.

The Lite version of the software, designed for C, C++, and Java source codes, was made to be deployed on the same PC where the code resides, making it a streamlined tool that is easy to use. The user just directs Predictive Lite to the target code, and the metrics are automatically applied. This simple interface also allows the user to



The Predictive product suite analyzes and predicts defects in software projects, allowing the user to identify the best metrics and thresholds to apply across one or more projects.

monitor progress as the tool performs its analysis. When the analysis is completed, all error-prone code is organized hierarchically. A user can typically expect to see a complete analysis within minutes, as Predictive Lite possesses an average processing speed of 1,500 lines of code per second (on a standard system with a 2.8 GHz processor and 512 MB of RAM).

In an example of application, Houston-based GB Tech, Inc., a technical services company providing engineering services to government, recently managed a small but critical software project for the U.S. Department of Defense (DOD), using Predictive Lite. The project involved software code, consisting of only 5,000 lines, that operated a battery management system for the DOD's Joint Strike Fighter manned airplane. Despite the diminutive number of lines, hazard criticality was rated extremely high. This code had to be tested extensively to ensure safety, given that the battery was a lithium ion battery, which can overcharge and possibly explode.

GB Tech first used a more expensive tool for structural code coverage before trying Predictive Lite. Later, it ran both tools on the same code to perform a comparison, which demonstrated that Predictive Lite produced consistent results and exceeded the abilities of the more expensive tool, in terms of architectural complexity and calculations.

Predictive Pro, the second product belonging to the Predictive portfolio, was launched in June 2005 and includes standard and heuristics modes, as well as a trend-identification mode that analyzes historical code to find patterns and trends. The standard mode is used early in the software life cycle, before error data are accumulated. As such data pile up, the heuristics mode can be employed to identify the error-prone code. This problem-solving mode uses an artificially intelligent engine to analyze the metrics and associated error data to learn what metric thresholds predict error within a specific project. (Integrated Software Metrics reports that results are very stable and more precise than those of domain experts

Earning 9 out of 10 stars from CRN magazine, Predictive Pro gives quality assurance managers "a tool with teeth." The software is deployed on the same PC where the targeted code resides. The user just directs Predictive Pro to the code and metrics are generated. The output shows every module in the project, the associated metrics, its location, and error data for that module. Predictive Pro's color-coding scheme of red, yellow, and green readily identifies the code modules of high, medium, and low risk, so that errors can be found easily.

who have a strong technical understanding of software analysis.) A user can then continue to use the heuristics mode through most of the software development cycle. Near the end of the cycle, the user can switch to the trend-identification mode to identify trends and find the chronic problems in the system.

The final product, Predictive Server, was released in September 2005. Predictive Server contains all of the features of Lite and Pro, and is scalable for network or distributed deployment. Unlike its predecessors, however, it was developed as a Web-based risk management tool for multiple software projects, and it facilitates collaboration among project managers, developers, and software quality assurance professionals. Essentially, the software aims to meet the demands of having to manage multiple software projects in a networked environment.

When Predictive Server is used on a software project, error data and metrics are compiled and stored in the software's knowledge database from the very beginning

of the project until the end. Thereafter, this historical database of metrics can be used for other software projects in the enterprise. Because Predictive Server is Web-based, all authorized software project teams can access and update the database, helping the organization deliver better software and save development costs.

In late 2005, Integrated Software Metrics announced a 200-percent surge in sales of its Predictive error-prediction tools. Driving this boost were new customers in the telecom, energy, technology, and government markets, including organizations such as Compagnie Financière Alcatel (Alcatel); Chevron Corporation; LogLogic, Inc.; and Northrop Grumman Corporation.

Spinning back to NASA, the Glenn and Goddard field centers are currently using the entire Predictive suite. Both centers rely on the products for critical code that supports NASA's Earth-orbiting spacecraft. ❖

Predictive™ is a trademark of Integrated Software Metrics, Inc.
Java™ is a trademark of Sun Microsystems, Inc.