

A Knowledge-based Geo-Spatial Decision Support System for Drought Assessment¹

Dustin Kozal, Matt Culver, Sherri Harms
University of Nebraska—Kearney
Kearney, NE 68849
{kozaldg, culvermg, harmssk}@unk.edu

Drought affects virtually all regions of the world and results in significant economic, social and environmental impacts. The Federal Emergency Management Agency estimates annual drought-related losses in the United States at \$6-8 billion, which is more than any other natural hazard. Congress enacted the Agricultural Risk Protection Act of 2000 to encourage the United States Department of Agriculture (USDA) Risk Management Agency (RMA) and farmers to be more proactive in managing drought risk. Current drought crisis management remains reactive, where action is dormant until the onset of threat. Despite billions of dollars in lost revenue, no comprehensive system exists for drought (Wilhite, 2000).

Through the National Science Foundation (NSF) Digital Government program, RMA is working with the University of Nebraska, the National Drought Mitigation Center, and the High Plains Regional Climate Center to develop geospatial decision support tools to address agricultural drought hazards and identify regions of vulnerability in the management of drought risk. The goal of the National Agricultural Decision Support System (NADSS) research project is to develop a support system of geospatial analyses that will enhance drought risk assessment and exposure analysis (Goddard et al., 2003).

The NADSS is a collection of decision support tools designed to help agricultural producers assess a variety of risks. The architecture is separated into four layers: knowledge, information, data, and presentation. The data layer contains unmodified data meant for transparent access, where as the information layer requests data from the data layer and develops it into information based on domain-specific servers. Examples of servers in this layer are data interpolation and map servers, which may be either domain independent (e.g., spline interpolation) or domain specific (e.g., terrain regression). The knowledge layer utilizes the information layer to develop domain-specific knowledge; it incorporates sequential data mining techniques, simulation modeling, and other knowledge analysis algorithms. Knowledge layer tools allow experts to extract meaningful interpretations for the complex problem of drought risk management. The intent is that decision-makers will interact with this layer via the user presentation interface to build and gather domain-specific knowledge.

The knowledge-based geo-spatial decision support system (DSS) for drought assessment is set within the knowledge layer of NADSS. This system provides real time assessment of drought associations with global climatic events in an easily accessible online interface. Two separate data mining algorithms represent the knowledge discovery in the project, the Representative Episodal Association Rules (REAR) (Harms et al., 2001) and Minimal Occurrences with Constraints and Time Lags (MOWCATL) algorithms (Harms et al., 2004). In previous research, these methods have been noted as effective tools for identifying relations between global climatic events, such as La Niña, and local drought conditions (Harms et al., 2003).

The knowledge discovery methods in the DSS (REAR and MOWCATL) are based on association rules that discover relationships between influential parameters. Association rules are easy to understand; most domain experts can relate to the expressed rules and use them in their decision making process. The

¹ This research was supported in part by NSF Digital Government Grant No. EIS-0091530 and NSF EPSCOR Grant No. EPS-0091900.

REAR algorithm finds episodes of events that occur together in a relatively short time interval, while MOWCATL finds patterns in one or more sequences that precede the occurrence of patterns in other sequences. These approaches are well suited for problems that have periodic occurrences when the signature of one sequence is present in another, even when multiple sequences are not globally correlated or spatially co-located.

The drought assessment DSS is functional over weather stations in the United States, with the ability to compute relationships and rules based on a range of local and global indices. The system begins when the user connects through the presentation layer to a dynamically updated webpage, where a series of parameters are selected. Using the built-in Common Object Request Broker Architecture (CORBA), data and execution requests are passed through back-end Java class and the resulting Java Beans. Communicating with the three lower layers in the NADSS architecture (knowledge, data, information), CORBA collects and sends the requested weather data. To simplify calls, input and output parameters are converted into XML strings by a customized XML parser. After the XML parser reassembles the settings and information from the requested weather data, the knowledge discovery algorithms are executed. Because the knowledge discovery methods were written in C++, the DSS uses Java Native Interface (JNI) to execute the methods as shared libraries. Upon its return, a separate C++ XML parser formats the result data. Finally, the back-end updates the DSS online interface with a table of results presented to the user's web browser.

Other tools or clients can use the knowledge discovery system without the web-based front-end due to the NADSS architecture and the DSS placement in the knowledge layer. NADSS allows each layer to make requests to any lower-level layer using a standard open interface. For example, the knowledge discovery tool in the knowledge layer may request a drought index for a given region from the information layer. If the data is not available (i.e., has not been pre-computed), the information layer will be able to retrieve the data from distributed spatial and relational databases making standard queries to the data layer to compute a response for the higher-level request. The NADSS server updates and processes new data frequently to remain up to date.

The web-based DSS will be an effective drought risk management tool for farmers, experts and general end-users. This type of DSS is critical to the future success of RMA and its responsiveness to the nation's farmers. With the effective use of this tool, RMA will have improved its efficiency and hold a greater impact in communicating drought risk management.

Over time, the system will be tweaked for optimal usage for both the client and server with performance enhancements include caching user queries and automating variable constraints. A case-based reasoning computational steering system is being developed that will guide the users in their choice of parameters. Additionally, GIS mapping is being integrated into the DSS to allow for extensive exploration into regional drought patterns.

References

- Goddard, S., Harms, S., Reichenbach, S., Tadesse, T., Waltman, W., "Geospatial Decision Support for Drought Risk Management," *Communications of the ACM*. 46 (1), January 2003, pp. 35-38.
- Harms, S.K., J. Deogun, J. Saquer, and T. Tadesse, "Discovering Representative Episodal Association Rules from Event Sequences using Frequent Closed Episode Sets and Event Constraints," *Proceedings of the 2001 IEEE International Conference on Data Mining (ICDM '01)*, San Jose, California, November 2001, pp. 603-606.
- Harms, S. K., Deogun, J., "Sequential Association Rule Mining with Time Lags", *Journal of Intelligent Information Systems (JIIS)*. 22 (1), January 2004, pp. 7-22.
- Harms, S. K., Deogun, J., Goddard, S., Building Knowledge Discovery into a Geo-spatial Decision Support System, *Proceedings of the 2003 ACM Symposium on Applied Computing*, Melbourne, FL, March 2003, pp. 445-449.
- Wilhite, D.A.: 2000, Drought as a natural hazard: concepts and definitions. In: D.A. Wilhite (ed), *Drought: A Global Assessment*, Natural Hazard and Disasters Series, Routledge Publishers, U.K.