Do not distribute. Submitted to *AMIA Annu Symp Proc*. Accepted June 2006.

Predicting Information Preferences from Information Logs

Mohamad El-Hajj, MSc; Robert S Hayward, MD, MPH; Tanya K Voth, MLIS
Centre for Health Evidence, University of Alberta, Edmonton, Alberta, Canada

Abstract*: This poster describes questions that can be answered about how clinical and knowledge-based resources are used in an integrated information environment that track how diverse applications are initiated, given focus, interacted with, and linked to other applications. Results show that it is possible to use decision trees to predict information patterns and preferences.*

Background: Health organizations are increasingly tasked with supporting clinicians' information needs at the point of clinical decision-making. Accordingly, most health institutions acquire and deploy diverse information resources alongside core clinical applications. Licensing costs and hosting methods vary depending on the type of system purchased (electronic texts and journals, synopses, guidelines, handouts, etc.) and are frequently managed by different departments. Without a comprehensive view of how all tools are used, it is difficult for administrators to ensure that the right applications are available at the right time for the right groups of users. Although others have asked clinicians about their needs and preferences,[1] self-reported data may not reflect what actually happens.

Decision trees support "intelligent" data mining techniques that can be used to predict choices when applied to sufficiently detailed records of information-seeking behaviors. Decision trees allow a set of input properties to be defined, and then use Boolean functions to predict patterns (within a pre-determined range of confidence) based on historically similar situations. This poster highlights how decision trees have helped answer important questions about information resource usage.

Methods: Data was collected using the VIVIDESK Internet "desktop" technology (www.vividesk.com) over a one year period (2005) for internal medicine practitioners (123 users) and trainees (90 users) at the University of Alberta. Data collection began the moment a user launched the desktop and ended when they closed the desktop or timed out. The desktop tracked which applications were in focus and allocated usage times accordingly; providing information about duration, order, context and method of use rather than just frequency. Data about clinical and knowledge-based applications were analyzed (98 applications in total). Knowledge resources were classified as studies, synopses, syntheses or systems.[2] Data was analyzed by application type, user type, frequency and duration of use, time and location of use and user interaction.

Results: Our results show that it is possible to use decision trees to answer practical questions that may support decisions about resource allocation in health organizations. Examples of the types of questions that can be answered include:

1. Does frequency and duration of use vary based on the time of the day?
2. What types of users use specific applications during certain time periods?
3. How often and for how long will different types of users use specific applications?
4. Which applications are used together, in what order?
5. Do application access services, such as single-sign-on, affect which applications are used, when and where?
6. Do application integration services, such as inter-application communications and context management (CCOW) affect which applications are used, when and where?
7. What are the peak usage times for specific applications and how many simultaneous licenses are required to support peak usage?
8. Which days of the week do users tend to use applications requiring the most helpdesk support?
9. Which applications are used, by whom and in what way, outside of normal work hours or locations?
10. Which applications evidence most intense mouse or keyboard activity, with what implications for hardware requirements?

Conclusion: Our research shows that clinician information needs vary and that prediction of usage patterns can address practical questions. The answers may guide development of more efficient clinical information environments.

References

1. Dawes M, Sampson U. Knowledge management in clinical practice: a systematic review of information seeking behavior in physicians. Int J Med Inform 2003; 71(1):9-15.
2. Haynes RB. Of studies, syntheses, synopses, and systems: the "4S" evolution of services for finding current best evidence. ACP J Club 2001 Mar-Apr; 134(2):A11-3.