

A secure electronic diary and data collection tool

Kenny Morrison & Ian W. Ricketts

School of Computing
University of Dundee, DD1 4HN

Martyn C. Jones

School of Nursing and Midwifery
University of Dundee, DD1 4HJ

Derek W. Johnston

School of Psychology
University of Aberdeen, AB24 2UB

Nigel B. Pitts

Dental Health Services Research Unit
School of Medicine
University of Dundee, DD2 4BF

Frank M. Sullivan

Clinical & Population Sciences & Education
School of Medicine
University of Dundee, DD2 4BF

Abstract—Handheld computerised data-collection tools have been shown to be more accurate, less susceptible to recall errors and to encourage higher user-compliance than pen and paper alternatives. Pocket Interview is a configurable electronic diary and data collection tool that provides the ability to design and deliver questionnaires and collect data securely using handheld devices and desktop computers. The client software can be set to prompt the data contributor, via audible prompts, to provide entries at fixed times and preset intervals. The device also allows the users to record personally significant incidents out with the schedule times. Pocket Interview is designed for use, after minimal training, by both the data-provider(s) and the administrator. An administrator can define and/or amend the data they would like to be collected, via a graphical user interface. The system has been shown to be usable by both administrators and users.

Keywords: *electronic diary, handheld computers, data collection, encryption*

I. INTRODUCTION

The Pocket Interview system provides an easily usable and configurable data collection tool that provides the ability to design and build questionnaires and securely collect data using handheld devices and desktop computers.

Mobile technologies are being used increasingly to gather data in clinical and other research studies. Traditionally, researchers have used pen and paper and subsequently entered the data into statistical software or subcontracted the task to a data entry company. The latter approach produces fewer errors but can be costly [1]. Direct data entry into a handheld computer can have several advantages compared to the traditional methods such as pen & paper. Capturing data using paper forms and then entering the results into a computer doubles the data entry required and increases the opportunity for introducing errors. Using mobile computers can save time by avoiding the double entry of data and reduce cost by reducing the need for skilled staff. Direct entry also offers the potential to improve the quality of the data, for example, by range checking values immediately they are entered by the

user. Additionally, researchers can analyse the results instantaneously. Accurate results are assured every time as most input is point and click, via menus and multiple choice questions, so data is captured quickly and accurately. This consistent input means that data can be collated easily and summary reports produced more quickly.

A. The use of handheld computers as electronic diaries

Pocket Interview can be used for a variety of data collection tasks although it was primarily designed to be used as a configurable electronic diary (ED). EDs allow researchers to collect a user's self-reported data and thereby to monitor states, events and behaviour. Traditional methods of self-report tend to use pen and paper. Clinical researchers first used paper diaries for self-reporting in the 1940's, latterly however, EDs have become increasingly popular [2].

B. Pen & Paper vs. Electronic

Comparisons of pen and paper with electronic alternatives have been the subject of recent research. Studies report more complete entries using EDs and generally far fewer errors. It has also been shown that users answers are more forthright and honest when using computer-assisted self-report [3] so EDs may be a more effective means of collecting sensitive information.

User entries can be date and time stamped so a precise record of when entries are made and how long entries take (even for individual questions) can be gathered. EDs are more efficient as less time is required to complete them and there are no problems or issues with deciphering hand written entries or losing paper forms. A wide variety of question formats can be made available and the requirement for a response can be enforced. Users can also be provided with examples to assist their understanding and shortened training times of five minutes have been demonstrated [4].

However, one area that has provoked recent debate is the issue of user-compliance with the data entry. This is the extent to which users comply with instructions, particularly with

regard to the timing of the diary entries. Stone, Shiffman *et al* published research in 2002/2003 that proved alarming for researchers using paper diaries [5]. They developed a novel paper diary fitted with a hidden photosensor that detected and recorded when the diary was opened and closed. The diary cards could not be removed and the users of the diary were not aware of the photosensor or the purpose of the study. The findings were that patients completed 90% of the entries but only 11% were actually completed within 30 minutes of the scheduled assessment times. They also found that 75% of users submitted entries for days that the diaries had not even been opened and that 45% of users actually forward-filled entries. Large batches of entries were even completed in the car park immediately before returning the diaries. It was suggested that users may have missed diary entries due to lack of motivation, forgetfulness, or lack of opportunity and that they didn't want the embarrassment of missed entries. It was suggested that perhaps users had not appreciated that they were actually defeating the purpose of the diaries.

A comparison group was run using EDs that prompted users with an audible signal and only allowed entries within a limited window of time and thereby prevented backfilling and forward-filling. Using EDs, 94% of diary entries were within 30 minutes of the scheduled times.

These studies stirred debate amongst researchers who argue that motivating and educating participants about the importance of compliance was a bigger factor than whether studies are administered using pen & paper or EDs [6].

C. Using Electronic Diaries to collect data in real-time

The use of EDs to record data contemporaneously, also termed real-time, and in the user's natural environment has previously been referred to as 'Ecological Momentary Assessment' [7] and 'Experience Sampling' [8]. Research has shown that a user's recall is unreliable and rife with inaccuracies and bias. People have also been shown to exhibit selective memory, especially for the most recent or the most severe incidents during a time period [9]. Gathering data in real-time has the advantage that users report before they've had the time to reflect and reinterpret their experiences after the fact.

Real-time data capture has been implemented in various ways:

- Personal Digital Assistants (PDAs) - currently the most widely used method.
- Pen and paper plus prompting from an electronic pager or programmable watch. The participant is asked to write down what they are doing as they are doing it which may be a burden and also affect the activity itself.
- Laboratory settings - it may be difficult to simulate realistic behaviour in a laboratory and may also fail to capture the influence of the setting on the participant's behaviours. It may also be unethical to conduct a laboratory study that induces stress in the user.

- Direct observation which can be costly, time-consuming and where the presence of an observer may affect the behaviour of the participant.

- Interactive voice response (IVR) - users provide data using touch-tone phones. IVR is often criticized as being unhelpful and difficult to use if poorly designed and therefore should be designed with the minimum of complexity.

- Internet - more recently researchers are using surveys which are administered over the web.

Using an ED can enable a variety of diary protocols to be employed. These can be time-based protocols such as fixed intervals where entries are requested at the same time every day or random intervals where the user is prompted at random throughout the day. The users will be interrupted during their normal daily activities that, if prompts are frequent, may prove irritating. Diaries can also use an event-based protocol where the user decides when to make the entries, normally after some noteworthy incident.

II. IMPLEMENTATION

After initial requirements gathering two user groups were identified: the administrators of the system and the data providers. Two subcomponents of the system were also identified: a PDA client used for prompting and data collection and an administration tool, run on a desktop PC, used for installing the client software, defining data collection schedules, defining the forms/questionnaires and collecting the data from the devices. Expert users offered their expertise during the requirements gathering and the development process therefore ensuring user-involvement throughout the software's life-cycle.

A. Client Software

The client software, used for data collection, is run on a Pocket PC handheld computer. The interface to the questionnaire is created dynamically by the client subcomponent. It reads the questions from a configuration file generated by the administration component. The software displays questionnaires, which can extend to multiple pages where necessary, and thus avoids the use of lengthy scrollbars. Text-entry is performed using an onscreen keyboard. Questionnaires can include 'back' and 'next' buttons allowing the user to review answers.

The software interface can be displayed in 'kiosk' mode. This is where the questionnaire fills the entire screen. This mode can make the device seem less intimidating for the inexperienced user, as the interface is less cluttered and it also prevents the user from accessing the other functionality of PDA. Kiosk mode can be switched on/off by the administrator.

When the system is active some entries may only be made when the user is prompted to make that entry. Included on the client is a 'demo' mode for demonstration and training users. Whilst in demo mode all entry types are visible on the main screen and data is not stored.

The software can prompt the user to make diary entries, using audible prompts/alarms that are triggered according to a schedule set by the administrator. The client will record whether the user has missed entries. If the data-providers do not respond to an alarm they can be reminded. A 'snooze' facility is also provided. There may be situations where the user may not want to be disturbed, so they may 'snooze' the device for up to one hour. Multiple protocols of prompts may be run simultaneously. As well as fixed schedules and random intervals, the device also supports spontaneous data entry that might be triggered by an external event occurring e.g. a post-emergency report.

The client includes a system area, which is password protected, and allows administrators to modify settings such as volume of the prompt/alarm and the type of sound used. The administrator can also edit the number and length of time (in minutes) between reminders.

The system allows the data collected to be encrypted using the RC4 cipher. This conforms to the recent NHS Scotland standard for data held on laptops, memory sticks and all other mobile devices [10]. The standard, introduced in September 2008, outlines the minimum requirements for the protection of mobile data within NHS Scotland. It states that all mobile data should be protected through the application of appropriate security controls and encryption regardless of the sensitivity of the information and was introduced in response to recent incidents that involved the discovery of unencrypted medical data on portable computing devices.

B. Administration software

The second component of the system is the administration tool that is run on a desktop or laptop PC and is organised into various sections.

- Client manager - used to install the client software on the Pocket PC together with the Microsoft .net compact framework.
- Schedule manager - consisting of a graphical user interface used to create the schedules of prompts. The administrator can add individual prompts or the prompts can be organised in to groups. The administrator can allocate the start and end times and also the number of alarms. The alarms will be evenly spaced. To prevent predictability and user-anticipation, an element of randomness can be added to individual alarms. Alarms can also be deleted or edited individually. Once the schedule has been finalised it is saved to file before copying to the data-collection device. The schedule manager also allows the administrator to edit schedules already on the client or data-collection device by copying them from the device to the administrator's PC, making the required changes and then copying them back to the client.
- Data manager – when the clients' devices have been returned from the data-providers the data will be retrieved. The data manager allows the administrator to collect and organise the data in a suitable manner. The data is in XML format when copied from the client to desktop. The data manager allows the data to be entered into a Microsoft Access database where it is easier to view and manipulate the data.

- Settings manager - allows the administrator to adjust the client settings and administer a master settings file. An administrator may be dealing with a number of clients and to avoid editing settings on multiple devices make any changes here and then copy the settings on to each client.

- Questionnaire manager - used for creating and editing existing questionnaires and consists of a graphical user interface (GUI). Questionnaires are stored as configuration files in XML format. XML can appear intimidating to the non-computer programmer and can be difficult to edit. Like many computer languages the syntax may appear fussy and small changes and elements incorrectly placed can cause malfunctions. Using the GUI, shields the user from the complexities of XML.

There are a wide range of question types available, including text boxes, drop-down combo boxes, track bars (visual analogue scales which can detect where the data-provider clicks), check boxes, Likert scales and radio buttons.

The questionnaires have a hierarchical structure. Sequences contain groups which contain questions. Groups of questions may appear in different sequences. The manager includes a preview facility that emulates the client software and allows the administrator to view questionnaires prior to copying to a PDA. The software includes user help in the form of documentation and videos covering the more complex parts of the system administration such as branching within questionnaires.

C. Desktop version

A version of Pocket Interview has been developed to run on a desktop or laptop PC. It may not always be necessary to provide users with a handheld computer as the data-provider may be located in front of a PC or has easy access to one. The Pocket Interview client will run in the background on the PC. Its presence will be indicated by an icon in the taskbar and a user will be prompted by a balloon appearing in the bottom right corner of the screen. The user clicks on the balloon to open the software and make an entry. Therefore the system will not appear intrusive. The interface looks very similar to the handheld version in terms of size and layout. Downloading of the questionnaires and schedules, and uploading of the contributed data can be done automatically if an Internet connection is available .

III. RESULTS AND DISCUSSION

There are PDA-based data collection tools freely available to the academic community [11] but administrating these systems is not always as straightforward as it could be. They tend to rely on manual editing of the configuration files. These tools may be restricted to older versions of the Palm and Windows CE operating systems and therefore may only operate on outdated hardware. Their protocol definition process may also require users to modify configuration text files. A process that can be susceptible to error and may require extended knowledge of the syntax. Pocket Interview uses XML to define the questionnaires and schedules which can appear quite daunting to the non-computing user. Therefore Pocket Interview is easily configurable through the use of a user-

friendly graphical user interfaces. This allows the administrator complete control and flexibility when designing and administering data collection. It is also believed Pocket Interview is the first configurable diary that is capable of running on PDAs, desktop computers and smart phones.

There are also some commercial applications available but these can be expensive and a separate license may need to be purchased for each data collection device.

A. Evaluations

There has already been great interest in the software from a broadening group of medical users. Evaluation has begun with several teams who wish to use the software to further their research and are defining and delivering their own questionnaires and schedules of prompts. Their interests include ambulatory methods of measuring work stress, factors influencing hand hygiene practice in nursing staff and monitoring the mental wellbeing of veterinary graduates. These teams are currently evaluating Pocket Interview and have demonstrated that they can use all aspects of the system with minimal assistance.

B. Future plans

More recently researchers have developed configurable smart phone based systems as a means of gathering data [12]. At present the Pocket Interview software is compatible with smart phones running the Windows Mobile Professional operating system and with some further development will run on all Windows-based mobiles. However a report in 2008 [13] estimated that Windows mobile only accounted for 3.5% of the global smart phone market. Symbian is the platform adopted by Nokia, Sony Ericsson and Panasonic amongst others and it accounts for 46.6% of the global market. The Apple iPhone has 17.3% and the Blackberry has 15.3%. Therefore work is underway to create a mobile phone version of Pocket Interview which is compatible with all mobile platforms.

There are many potential advantages of using mobile phones. These are devices that many users already own and are comfortable operating. Mobile phone users are more likely to keep the devices with them and are less likely to lose or forget them. There is also the potential for project savings as additional equipment may not be required.

As the Pocket Interview software evolves, the advantages of adding more functionality may become apparent. The ability to connect wireless Bluetooth devices such as accelerometers and to record audio will be integrated. Voice recording is also being introduced to support the users responding to open-ended questions. These often require longer answers so instead of making a lengthy text entry the user could dictate their response. The use of mobile phone networks to upload data will be added. Data may not need to be stored on the device as it could immediately be uploaded after an entry has been made. The use of global positioning system could also be explored and integrated.

Electronic diaries can be used to collect a user's self-reported data and to record data in real-time and in the user's natural environment. Pocket Interview provides the ability to design and deliver secure data collection using handheld devices and desktop computers. Through the use of an easy to use software tool, administrators can quickly build, test and distribute data forms. The responses subsequently provided by users can then be collected, organised and analysed in a timely fashion.

There are potentially many health-based applications and to encourage its use the software is available, at no cost, to the research community via the project website at: 'www.computing.dundee.ac.uk/acprojects/PocketInterview/'

V. ACKNOWLEDGEMENTS

The development of Pocket Interview has been funded in part by the UK Medical Research Council, the Social Dimensions of Health Institute (Universities of Dundee and St. Andrews), the Dental Health Services Research Unit and the School of Computing (University of Dundee).

- [1] B.A. Weber and B.L. Roberts, "Data collection using handheld computers", *Nursing research*, 2000. 49(3).
- [2] M.R. Hufford and A.L. Shields, "Electronic diaries: applications and what works in the field", *Appl Clin Trials*, 2002; p. 46-59.
- [3] R. Tourangeau and T.W. Smith, "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context", Oxford University Press on behalf of the American Association for Public Opinion Research, 1996, p. 275-304.
- [4] M. Palmblad and B. Tiplady, "Electronic diaries and questionnaires: Designing user interfaces that are easy for all patients to use", *Quality of Life Research*, 2004. 13(7): p. 1199-1207.
- [5] A.A. Stone, S. Shiffman, J.E. Schwartz, J.E. Broderick and M.R. Hufford, "Patient compliance with paper and electronic diaries", *Controlled Clinical Trials*, 2003. 24(2): p. 182-99.
- [6] A.S. Green, E. Rafaeli, N. Bolger, P.E. Shrout and H.T. Reis, "Paper or Plastic? Data Equivalence in Paper and Electronic Diaries", *Psychological Methods*, 2006. 11(1): p. 87-105.
- [7] A.A. Stone and S. Shiffman, "Ecological momentary assessment (EMA) in behavioural medicine", *Annals of Behavioral Medicine*, 1994. 16: p. 199-202.
- [8] R. Larson and M. Csikszentmihalyi, "The experience sampling method", *New Directions for Methodology of Social and Behavioral Science*, 1983. 15: p. 41-56.
- [9] N. Bolger, A. Davis, and E. Rafaeli, "Diary Methods: Capturing Life as it is Lived", *Annu. Rev. Psychol.*, 2003. 54: p. 579-616.
- [10] NHSScotland, "Mobile Data Protection Standard v1.0. 2008", Available from: <http://www.ehealth.scot.nhs.uk/wp-content/documents/mobile-data-protection-standard-pdf3.pdf>.
- [11] T. Conner, "Experience Sampling Resource Page," 2006, Available from: <http://psychiatry.uhc.edu/faculty/files/conner/ESM.htm>.
- [12] J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, J. A. Landay, "MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones" *MobiSys*, 2007.
- [13] Canalys, "Global smart phone shipments rise 28%", 2008, Available from: <http://www.canalys.com/pr/2008/r2008112.htm>.