APRIMORAMENTO DO SISTEMA DE *LOGS* DE INTERAÇÃO DE DISPOSITIVO TANGÍVEL MULTISSENSORIAL PARA SUPORTE À ALFABETIZAÇÃO

ANTONIO GABRIEL MAIA FICKEL[1]; EDUARDA PEREIRA MEDEIROS[1]; LAURA QUEVEDO JURGINA[1]; LEOMAR SOARES DA ROSA JÚNIOR[1]; MARILTON SANCHOTENE DE AGUIAR[1]

[1]Universidade Federal de Pelotas –
{agmfickel,epmedeiros,lqjurgina,leomarjr,marilton}@inf.ufpel.edu.br

1. INTRODUÇÃO

Estudos sobre o uso de interfaces de usuário tangíveis (TUIs) na alfabetização infantil indicam que a manipulação de letras físicas, associada a um processo multissensorial, favorece o aprendizado, sobretudo entre crianças com dificuldades de aprendizagem (Fan *et al.*, 2017; Polat *et al.*, 2019). Para auxiliar no aprimoramento dessas soluções, assim como na compreensão do comportamento dos usuários, a implementação de armazenamento automático de *logs* (registros) das interações com o dispositivo torna-se uma prática importante (Schneider e Blikstein, 2015). A utilidade desses dados para análise, porém, depende diretamente da precisão dos registros.

O Alfaba, base deste trabalho, é um dispositivo multissensorial voltado à alfabetização de crianças (Jurgina *et al.*, 2023). Sua interface principal é composta por oito baias destinadas à inserção de letras físicas, que o usuário deve preencher para formar a palavra solicitada visual e sonoramente pelo aparelho. Ao longo das sessões, o dispositivo oferece retorno imediato às tentativas por meio da iluminação das letras, ao mesmo tempo em que registra os dados da interação.

A primeira versão do sistema de *logs* de interação surgiu durante o desenvolvimento do primeiro protótipo do Alfaba, com o objetivo principal de monitorar tecnicamente o funcionamento do dispositivo. No entanto, o formato mostrou-se impreciso para capturar detalhes do comportamento dos usuários, o que tornou necessária a reformulação do sistema para viabilizar o rastreio de evidências dos casos de uso. Este trabalho apresenta essa mudança e analisa os seus efeitos sobre a qualidade dos dados coletados.

2. METODOLOGIA

Na versão 1 dos *logs*, cada linha reunia cinco colunas separadas por barra vertical: tempo de execução em milissegundos, letras inseridas, palavra proposta, tamanho da palavra e máscara de espelhamento (com “1” sinalizando uma letra inserida com a orientação incorreta). Um exemplo hipotético desse formato, referente à formação da palavra "casa", aparece na Tabela 1. As colunas B1 a B8 indicam a letra inserida em cada baia no instante da gravação; o caractere "Ƨ" representa a letra "S" inserida com orientação invertida.

Tabela 1 – Exemplo hipotético da versão 1 dos *logs*

Linha	B1	B2	B3	B4	B5	B6	B7	B8	Registro
1									*46822||casa|4|00000000*

Linha	B1	B2	B3	B4	B5	B6	B7	B8	Registro
2									*49878||casa|4|00000000*
3	C								*52934|c|casa|4|00000000*
4	C								*55990|c|casa|4|00000000*
5	C		Ƨ						*59046|cs|casa|4|00100000*
6	C	A	S						*62102|cas|casa|4|00000000*
7	C	A	S	A					*6265158|casa|casa|4|00000000*
8	C	A	S	A					*68214|casa|casa|4|00000000*

Foram identificados quatro problemas principais na versão 1, mitigados na versão 2:

1. **A gravação ocorria de forma periódica** (aproximadamente a cada 3 s), mesmo com o Alfaba ocioso, comportamento observável nas linhas 1 e 2 e nas linhas 3 e 4 da Tabela 1. Na versão 2, as escritas no cartão de memória ocorrem apenas mediante interação do usuário.
2. **Espaços vazios na formação da palavra não eram indicados**. Na linha 5, o log aponta “cs”, o que sugere ocupação das duas primeiras baias. Essa informação está incorreta e contradiz a máscara de espelhamento. A versão 2 resolveu isso ao representar a baia vazia com o caractere “_”: “c_s_____”.
3. **Não havia identificação de sessão** e a separação entre um uso e outro só podia ser inferida a partir da reinicialização do tempo de execução. Isso dificultava a análise por usuário. Na versão 2, acrescentou-se um código de sessão de quatro caracteres, sorteado e apresentado na tela quando o Alfaba é ligado (exemplo: “A1B2”). Como o dispositivo não dispõe de relógio de tempo real, é esse código que viabiliza a análise por usuário: basta anotar em registro externo o código e a quem ele corresponde e reiniciar o equipamento a cada troca de usuário.
4. **Ocorriam corrupções ocasionais dos dados**, como exemplificado na linha 7, cujo tempo de execução aparece corrompido, representando um salto incoerente com a linha seguinte. O problema foi resolvido utilizando um *mutex* para acessos ao cartão de memória, que é utilizado por mais de uma *thread* do Alfaba.

Os dados utilizados nesta análise foram coletados *offline* em testes com crianças, mediante assinatura do Termo de Consentimento Livre e Esclarecido pelos responsáveis e do Termo de Assentimento Livre e Esclarecido pelas próprias crianças. Para verificar empiricamente o efeito das mudanças descritas, que incluíram também a adição de cinco colunas a mais, comparou-se um arquivo da versão 1 a um da versão 2, gerados em épocas e a partir de interações diferentes.

3. RESULTADOS E DISCUSSÃO

A Tabela 2 resume as métricas de comparação. Considerou-se redundante toda linha idêntica à imediatamente anterior, dentro de uma mesma sessão, exceto pelas colunas de tempo. A redução dessas linhas evidencia a otimização: na versão 1, 88,02% das linhas documentavam estado inalterado, percentual zerado na versão 2. A adição de cinco colunas novas contribuiu para aumentar a mediana de bytes por linha em 96,6%, mas o custo de armazenamento se manteve praticamente o mesmo, com redução de cerca de 4,6% na mediana de bytes por minuto, já que a nova versão reduziu em 54% a mediana de linhas por minuto. Assim, a reformulação não alcançou economia significativa de espaço no cartão de memória, mas sim um aumento significativo na densidade da informação armazenada.

Tabela 2 – Comparação das versões 1 e 2 dos *logs*

Medida	**Versão 1**	**Versão 2**
Linhas analisadas	5.928	12.902
Bytes analisados	169.835	737.969
Linhas redundantes	88,02%	0%
Linhas defeituosas (formatação quebrada ou tempo implausível)	26	0
Mediana de bytes por linha	29	57
Mediana de linhas por minuto	20,68	9,50
Mediana de bytes por minuto	558,72	533,21

Quanto à integridade dos dados, comprometida na versão 1 pela ausência de tratamento adequado de *threads*, as 26 linhas defeituosas (0,44% do total) se dividem em dois tipos de falha: 5 apresentaram falha de leitura no formato definido e 21 apresentaram tempo de execução implausível, identificado quando o valor rompe a sequência temporal. Destas últimas, 12 trazem um valor inflado, sendo que em 11 delas o valor correto é recuperável ao se removerem os dígitos iniciais provenientes da linha anterior. As 9 linhas implausíveis restantes perderam dígitos. Na versão 2, o uso do *mutex* extinguiu as linhas defeituosas detectáveis.

Além das melhorias nessas métricas, a adição do caractere "_" para representar espaços vazios eliminou a ambiguidade da versão anterior, em que letras inseridas fora da sequência da palavra apareciam justapostas. O código de sessão também teve efeito positivo na análise dos dados pela equipe pesquisadora, pois permitiu a identificação direta das sessões individuais e dos usuários associados a elas (com registro externo).

4. CONCLUSÕES

A reformulação apresentada tornou o sistema de *logs* do Alfaba mais preciso como plataforma de coleta de dados de interação. Esse resultado apoia-se nas decisões de gravar somente após uma interação, tornar explícito no formato aquilo que antes precisava ser inferido, como a posição de cada letra na palavra e a fronteira entre um uso e outro, e controlar por exclusão mútua o acesso ao cartão de memória.

Todavia, a análise apresenta algumas limitações. A comparação é realizada entre arquivos com interações diferentes, de modo que as medianas por minuto podem refletir também intensidade de uso; a detecção de corrupção pode não perceber linhas defeituosas cujos valores sejam plausíveis; o código de sessão do formato novo não identifica o usuário sem o auxílio de um registro externo. O armazenamento também continua sendo feito em um único arquivo contínuo, o que dificulta o manuseio dos dados, problema que a implementação de uma terceira geração do sistema de *logs* pretende resolver, ao apresentar informações mais seletas em arquivos separados por sessão e por usuário.

Agradece-se ao CNPq, à FAPERGS e à CAPES pelo fomento à pesquisa que viabilizou este trabalho.

5. REFERÊNCIAS BIBLIOGRÁFICAS

FAN, M.; ANTLE, A.N.; HOSKYN, M.; NEUSTAEDTER, C.; CRAMER, E.S. Why tangibility matters: a design case study of at-risk children learning to read and spell. In: **CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS**, 35., Denver, 2017. **Proceedings...** New York: ACM, 2017. p.1805-1816.

JURGINA, L.Q.; AQUINI, L.G.; SOARES, R.I.; ROSA JÚNIOR, L.S. da; AGUIAR, M.S. de; PRIMO, T.T. Alfaba: A Tangible Solution to Support Brazilian Dyslexic Students in their Literacy Process. In: **IEEE GLOBAL ENGINEERING EDUCATION CONFERENCE - EDUCON**, 14., Kuwait, 2023, **Proceedings...** [S.l.]: IEEE, 2023. p.1-9.

POLAT, E.; CAGILTAY, K.; KARASU, N. Tangible objects and mobile technology: interactive learning environments for students with learning disabilities. In: ZHANG, Y.A.; CRISTOL, D. (Ed.) **Handbook of mobile teaching and learning**. 2.ed. Singapore: Springer, 2019. p.635-654.

SCHNEIDER, B.; BLIKSTEIN, P. Unraveling students' interaction around a tangible interface using multimodal learning analytics. **Journal of Educational Data Mining**, [S.l.], v.7, n.3, p.89-116, 2015.