

ABSTRACT

Thesis: 93 p., 10 fig., 7 tabl., 5 append., 19 sources.

CARDINALITY, UNIQUE EVENTS, BIG DATA, HASH FUNCTIONS, DATA PROTECTION

Object of research: statistical data on the events of the visits to the site and the purchase of goods.

Subject of research: application of methods for estimating the number of unique events for the assessment of these events.

Purpose: analyze the subject of research and to investigate the effectiveness of the application of algorithms.

Methods: algorithms for estimating the number of unique events, estimating their accuracy on the data taking into account the specifics of the data themselves.

Relevance: estimating the number of unique events in large-scale samples is extremely relevant for medical research and information security. Therefore, there is a need for the investigated methods and the search for optimal solutions in the search for the uniqueness of events

Results of the work - the uniqueness of the events was calculated on 132 samples by classical methods of counting unique events, and the result was improved after the initial processing of data

Ways of further development of the subject of research - further research of alternative methods of probabilistic calculation, expansion of used hash functions for more topical problems in modern medicine