#19

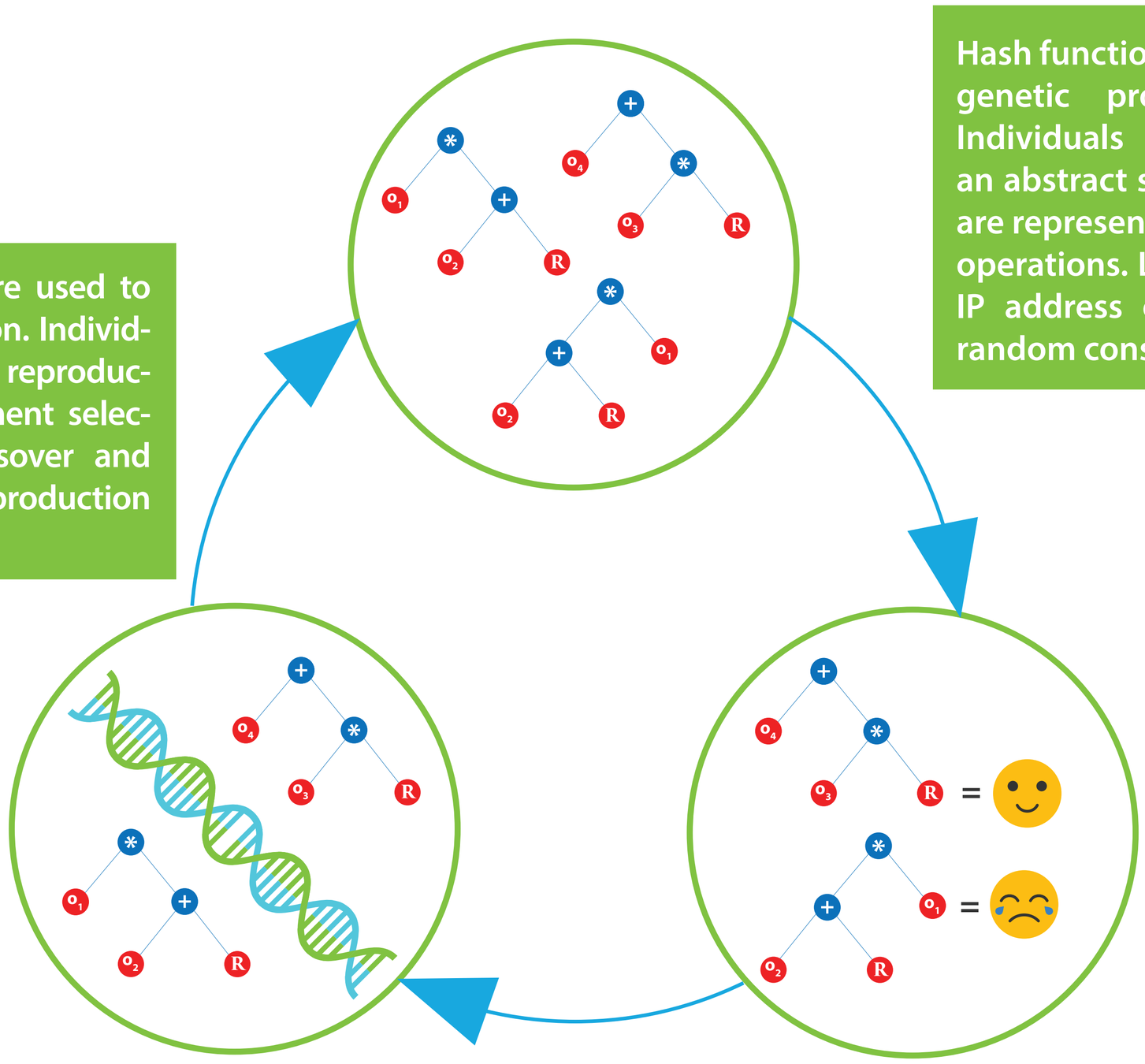
Marek Kidoň
xkidon00@stud.fit.vutbr.cz

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

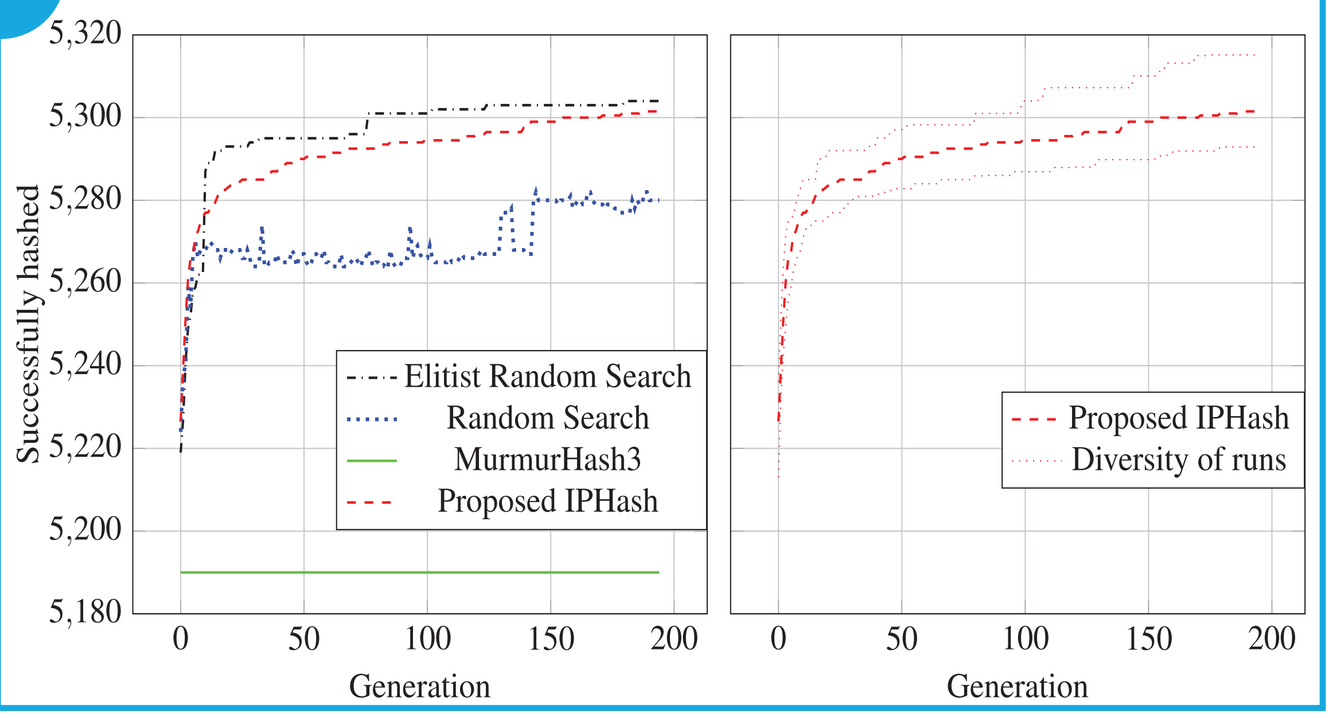
Excel@FIT 2016
http://excel.fit.vutbr.cz

Evolutionary design of domain specific non-cryptographic hash functions

Hash functions are evolved using the genetic programming algorithm. Individuals are represented by an abstract syntax tree where nodes are represented by common hashing operations. Leafs may represent and IP address octet or an ephemeral random constant.



1.



2.

Terminal set	Function set	Average best
$\{o_0..o_3, \Re\}$	$\{*, +, \wedge, \ggg\}$	5307
$\{o_0..o_3\}$	$\{*, +, \wedge, \ggg\}$	5291
$\{o_0..o_3, \Re\}$	$\{*, +, \wedge, \neg, \ggg\}$	5300
$\{o_0..o_3, \Re\}$	$\{*, \wedge, \ggg\}$	5306
$\{o_0..o_3, \Re\}$	$\{+, \wedge, \ggg\}$	5290
$\{o_0..o_3, \Re\}$	$\{*, +, \wedge\}$	5298
$\{o_0..o_3, \Re\}$	$\{+, \wedge\}$	5065

3.

$$f_{o_0,o_1,o_2,o_3} = (((\Re \wedge (o_1 \wedge o_3)) + (o_2 * \Re)) + ((\Re \ggg o_1) \ggg (((o_1 * o_0) \ggg o_1) * (\Re \ggg o_1))))$$

1. Comparison of the proposed IPHash genetic programming algorithm with random search algorithms and the human-created MurmurHash and diversity of the experiment. The result is an average of 50 independent runs.
2. The relation of various terminal and non-terminal sets to the resulting average (50 runs) best fitness. If the multiplication (*), addition (+) and rotation (>>>) are preserved, evolved hash functions are well performing.
3. Example of an evolved hash function. Such function is small in terms of operation count but well performing on the given dataset.