Problem 1: Network intrusion detection

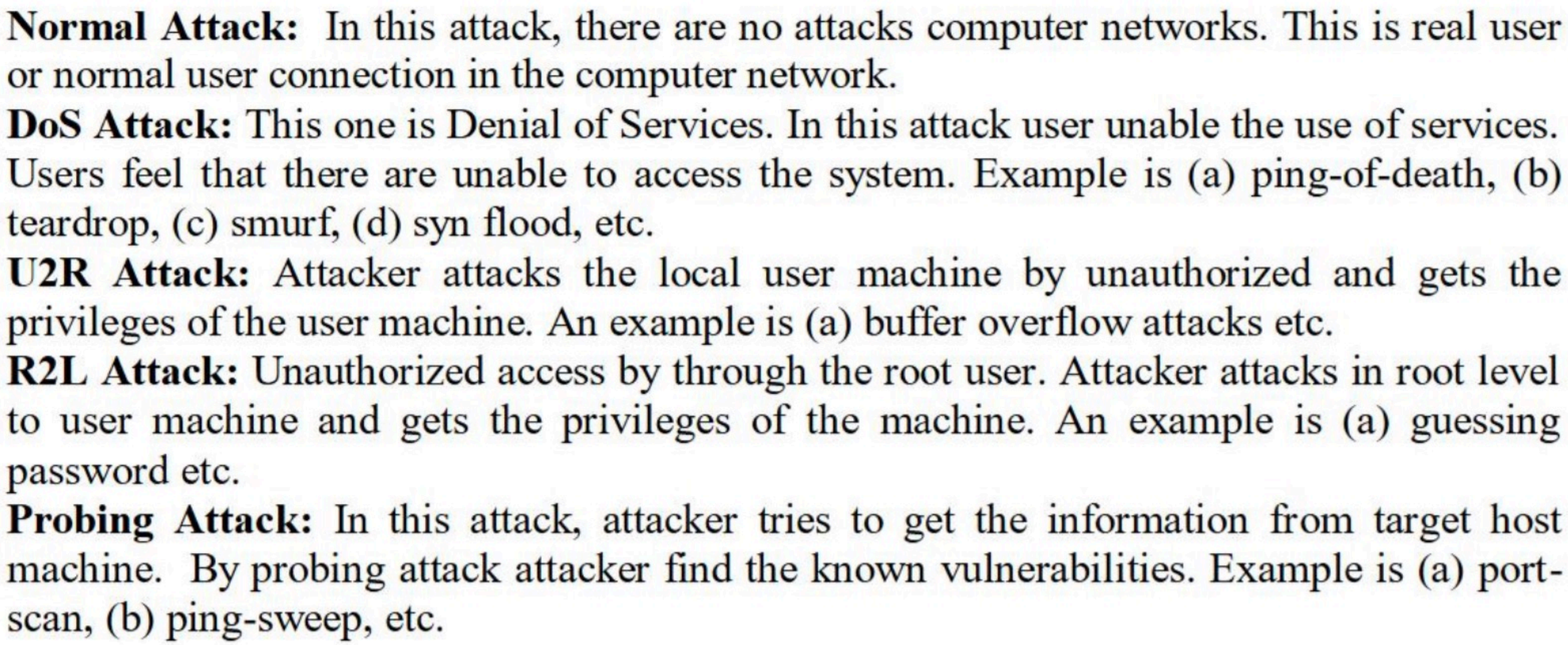
- Problem: use the networking features to analyze if it is an attack or normal packet? If it is attack, then classify it into a known type.

- Objective:
  - Create a machine learning model to correctly predict the class of the network data.
  - Use different preprocessing methods to clean, deduplicate and standardize the data.
  - Explore different types of feature selection algorithms, perform a comparative study to find the most effective feature selection algorithm.
  - Test the model with k-fold cross validation and check for any discrepancies.

- Data:
  - Tabular data, with ~400,000 samples.

- Paper: https://www.ijamtes.org/gallery/29%20conf-cse.pdf

Problem 1: Network intrusion detection



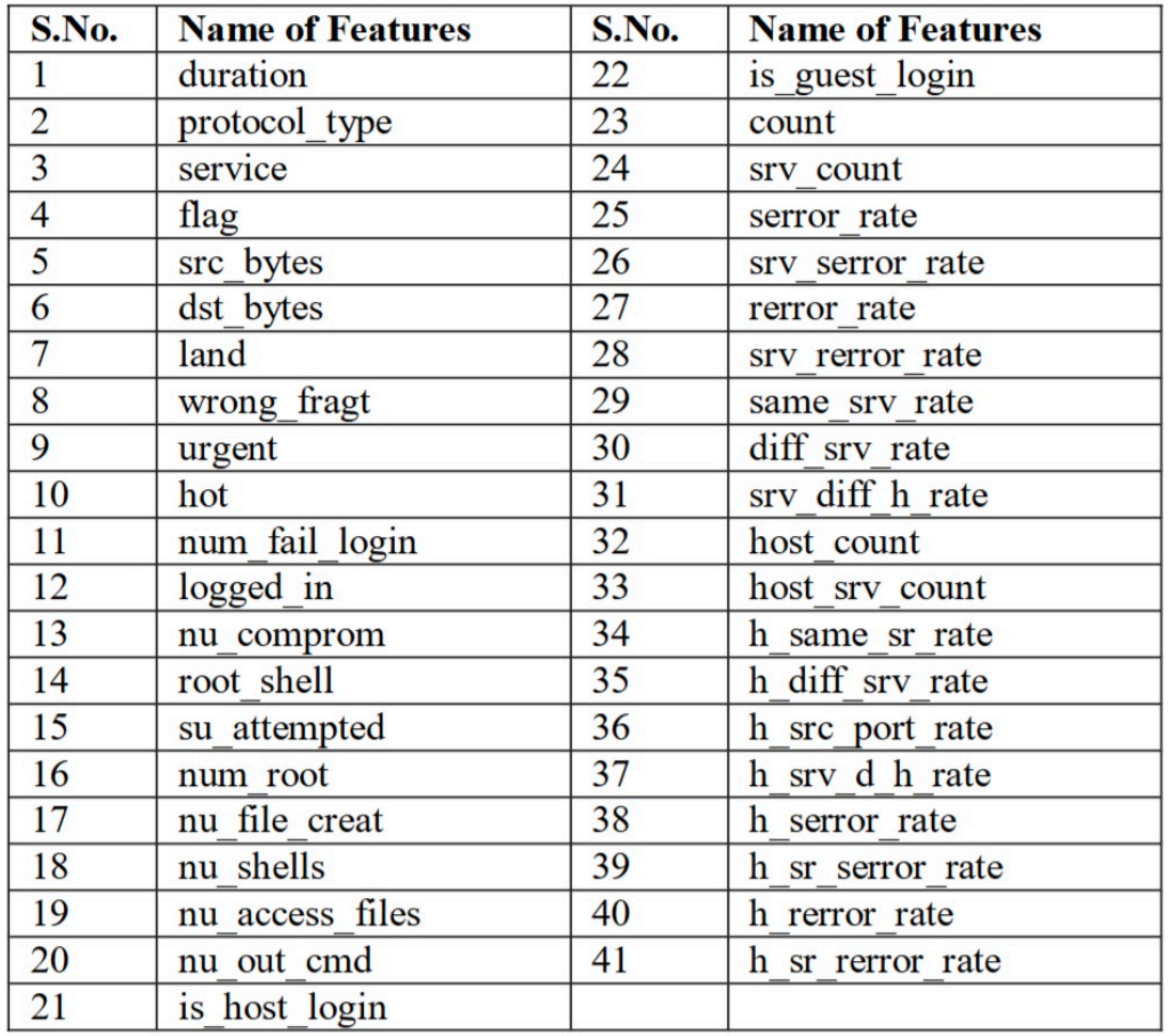
Normal Attack: In this attack, there are no attacks computer networks. This is real user or normal user connection in the computer network.

DoS Attack: This one is Denial of Services. In this attack user unable the use of services. Users feel that there are unable to access the system. Example is (a) ping-of-death, (b) teardrop, (c) smurf, (d) syn flood, etc.

U2R Attack: Attacker attacks the local user machine by unauthorized and gets the privileges of the user machine. An example is (a) buffer overflow attacks etc.

R2L Attack: Unauthorized access by through the root user. Attacker attacks in root level to user machine and gets the privileges of the machine. An example is (a) guessing password etc.

Probing Attack: In this attack, attacker tries to get the information from target host machine. By probing attack attacker find the known vulnerabilities. Example is (a) port-scan, (b) ping-sweep, etc.



S.No.	Name of Features	S.No.	Name of Features
1	duration	22	is_guest_login
2	protocol_type	23	count
3	service	24	srv_count
4	flag	25	serror_rate
5	src_bytes	26	srv_serror_rate
6	dst_bytes	27	rerror_rate
7	land	28	srv_rerror_rate
8	wrong_fragt	29	same_srv_rate
9	urgent	30	diff_srv_rate
10	hot	31	srv_diff_h_rate
11	num_fail_login	32	host_count
12	logged_in	33	host_srv_count
13	nu_comprom	34	h_same_sr_rate
14	root_shell	35	h_diff_srv_rate
15	su_attempted	36	h_src_port_rate
16	num_root	37	h_srv_d_h_rate
17	nu_file_creat	38	h_serror_rate
18	nu_shells	39	h_sr_serror_rate
19	nu_access_files	40	h_rerror_rate
20	nu_out_cmd	41	h_sr_rerror_rate
21	is_host_login		