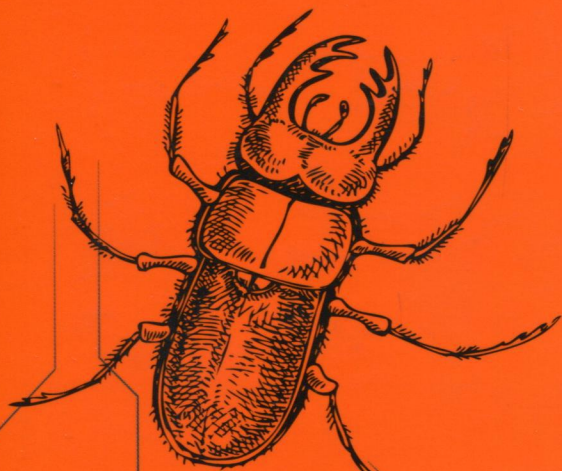


ВЛАДСТОН ФЕРРЕЙРА ФИЛО, МОТО ПИКТЕТ

ТЕОРЕТИЧЕСКИЙ МИНИМУМ ПО

COMPUTER SCIENCE



СЕТИ, КРИПТОГРАФИЯ
И DATA SCIENCE



ВЛАДСТОН ФЕРРЕЙРА ФИЛО, МОТО ПИКТЕТ

**ТЕОРЕТИЧЕСКИЙ МИНИМУМ ПО
COMPUTER SCIENCE**

**СЕТИ, КРИПТОГРАФИЯ
И DATA SCIENCE**



Санкт-Петербург • Москва • Минск

2022

ББК 32.973.23-018
УДК 004.3
Ф54

Феррейра Фило Владстон, Пиктет Мото

Ф54 Теоретический минимум по Computer Science. Сети, криптография и data science. — СПб.: Питер, 2022. — 288 с.: ил. — (Серия «Библиотека программиста»).

ISBN 978-5-4461-2945-4

Хватит тратить время на занудные учебники! Это краткое и простое руководство предназначено для читателей, не заботящихся об академических формальностях.

Большинство технологических прорывов нашей эпохи происходят в цифровой среде, создаваемой программистами. Ученые-компьютерщики объединяют различные области исследований и расширяют возможности этого нового мира. Чтобы научиться плавать в океане информации, необходимо разбираться в основах сетевых технологий, криптографии и науке о данных.

Вы узнаете, как эффективно манипулировать данными, освоите машинное обучение и современные концепции безопасности.

Раскройте мощь Computer Science и станьте гуру цифровой эпохи!

16+ (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

ББК 32.973.23-018
УДК 004.3

Права на издание получены по соглашению с Code Energy LLC. Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Информация, содержащаяся в данной книге, получена из источников, рассматриваемых издательством как надежные. Тем не менее, имея в виду возможные человеческие или технические ошибки, издательство не может гарантировать абсолютную точность и полноту приводимых сведений и не несет ответственности за возможные ошибки, связанные с использованием книги. Издательство не несет ответственности за доступность материалов, ссылки на которые вы можете найти в этой книге. На момент подготовки книги к изданию все ссылки на интернет-ресурсы были действующими.

ISBN 780-997316032 англ.

ISBN 978-5-4461-2945-4

© Computer Science Unleashed, Wladston Ferreira Filho and Raimondo Pictet, 2021

© Перевод на русский язык ООО «Прогресс книга», 2022

© Издание на русском языке, оформление ООО «Прогресс книга», 2022

© Серия «Библиотека программиста», 2022

Оглавление

ПРЕДИСЛОВИЕ	13
Так для кого эта книга	14
От издательства	14
Благодарности.	15
ГЛАВА 1. СВЯЗИ	16
1.1. Канальный уровень	17
Общие связи	18
MAC-адресация.	21
Кадры	24
1.2. Межсетевой уровень.	25
Межсетевое взаимодействие.	28
Маршрутизация.	28
Адресация местоположения	30
Интернет-протокол	31
1.3. IP-адресация	33
IANA	35
Провайдеры интернет-услуг	37
1.4. IP-маршрутизация	39
Таблицы адресов.	40
Точки обмена интернет-трафиком.	43
Интернет-магистраль	44
Динамическая маршрутизация	44
Петля маршрутизации.	45
Диагностика.	47
1.5. Транспортный уровень.	50
Протокол пользовательских дейтаграмм	50
Протокол управления передачей данных	53

Сегменты TCP	54
TCP-соединение	57
TCP-сокеты	59
Резюме	60
Дополнительная информация.	62
ГЛАВА 2. ОБМЕН ДАННЫМИ	63
2.1. Имена.	64
Домены	64
ICANN	65
Серверы имен	66
Запрос	67
Рекурсивный запрос	69
Типы записей	70
Обратный DNS-запрос.	72
Регистрация домена	73
2.2. Время.	75
Координированное универсальное время	76
Протокол сетевого времени	78
Серверы времени.	80
2.3. Доступ	82
Терминалы.	82
Telnet	85
2.4. Почта	86
Почтовые серверы	87
Simple Mail Transfer Protocol	88
Отправка электронных писем.	91
Получение электронных писем.	93
2.5. Сеть.	94
Язык разметки гипертекста.	95
URL-адрес	97
Протокол передачи гипертекста	99
Веб-приложения	101

Резюме	103
Номера портов	103
Одноранговое соединение	104
Безопасность	104
Дополнительная информация.	105
ГЛАВА 3. БЕЗОПАСНОСТЬ	106
3.1. Устаревшие шифры	107
Зигзагообразный шифр	108
Шифр подстановки.	109
Продукционные шифры.	111
Шифр Виженера	112
Шифр Вернама	113
Шифровальные машины	115
3.2. Симметричные шифры.	116
Потоковые шифры	117
Блочные шифры	120
3.3. Асимметричные шифры	124
Обмен ключами Диффи — Хеллмана	125
Шифры с открытым ключом	126
Цифровые подписи.	127
Цифровые сертификаты	128
3.4. Хеширование	129
Обнаружение злонамеренных изменений	130
Код аутентификации сообщения	131
Обработка паролей.	132
Доказательство существования	134
Подтверждение работы.	135
Небезопасные хеш-функции	136
3.5. Протоколы.	137
Безопасный доступ.	138
Безопасная передача	139
Другие протоколы	140

3.6. Хакинг	141
Социальная инженерия	142
Уязвимости программного обеспечения	145
Эксплойты	148
Цифровая война	150
Чек-лист защиты	152
Резюме	153
Дополнительная информация	154
ГЛАВА 4. АНАЛИЗ ДАННЫХ.	155
Сбор данных.	156
4.1. Сбор.	158
Виды данных	158
Получение данных	159
Ошибка выборки	161
4.2. Обработка	162
Первичная очистка данных	162
Анонимизация данных	167
Воспроизводимость	168
4.3. Обобщение	170
Количество	170
Средние значения	170
Изменчивость	172
Сводка пяти чисел	173
Категориальное обобщение	176
Корреляционная матрица.	176
4.4. Визуализация	179
Ящик с усами	180
Гистограммы	182
Точечные диаграммы	186
Временные ряды	190
Карты.	194
4.5. Тестирование	195
Гипотезы.	195

Эксперименты198
Р-значения200
Доверительные интервалы202
Резюме203
Дополнительная информация.205
ГЛАВА 5. МАШИННОЕ ОБУЧЕНИЕ206
Модели.207
5.1. Признаки.210
Адаптация данных211
Объединение данных216
Пропущенные значения.218
Утечка данных221
5.2. Оценка223
Оценка регрессоров225
5.3. Проверка работоспособности227
K-folds229
Монте-Карло231
Исключение по одному (leave-one-out).232
Интерпретация232
5.4. Подстройка233
Подстановка.235
Выбросы235
Нормализация236
Логарифмическое преобразование237
Биннинг238
Кластеризация238
Извлечение признаков239
Отбор признаков242
И снова утечка данных243
Выбор модели244
Заключительные шаги.245
Резюме246
Дополнительная информация.248

ЗАКЛЮЧЕНИЕ.249
БОНУСНАЯ ГЛАВА 6. ШАБЛОНЫ251
6.1. Соответствие253
Точка253
Множество.254
Обратное множество.256
Специальные символы257
6.2. Квантификаторы258
Фигурные скобки.258
Вопрос259
Плюс260
Звездочка260
Жадность261
6.3. Привязки.262
Каретка.262
Доллар.262
Граница263
6.4. Группы264
Захват групп.265
Чередование266
Резюме267
Дополнительная информация.269
ПРИЛОЖЕНИЯ270
I. Основания систем счисления270
II. Взлом шифра сдвига.271
III. Взлом шифра подстановки272
IV. Оценка классификаторов274
Компромисс в классификации279
Кривые ROC280
Многоклассовая классификация.282