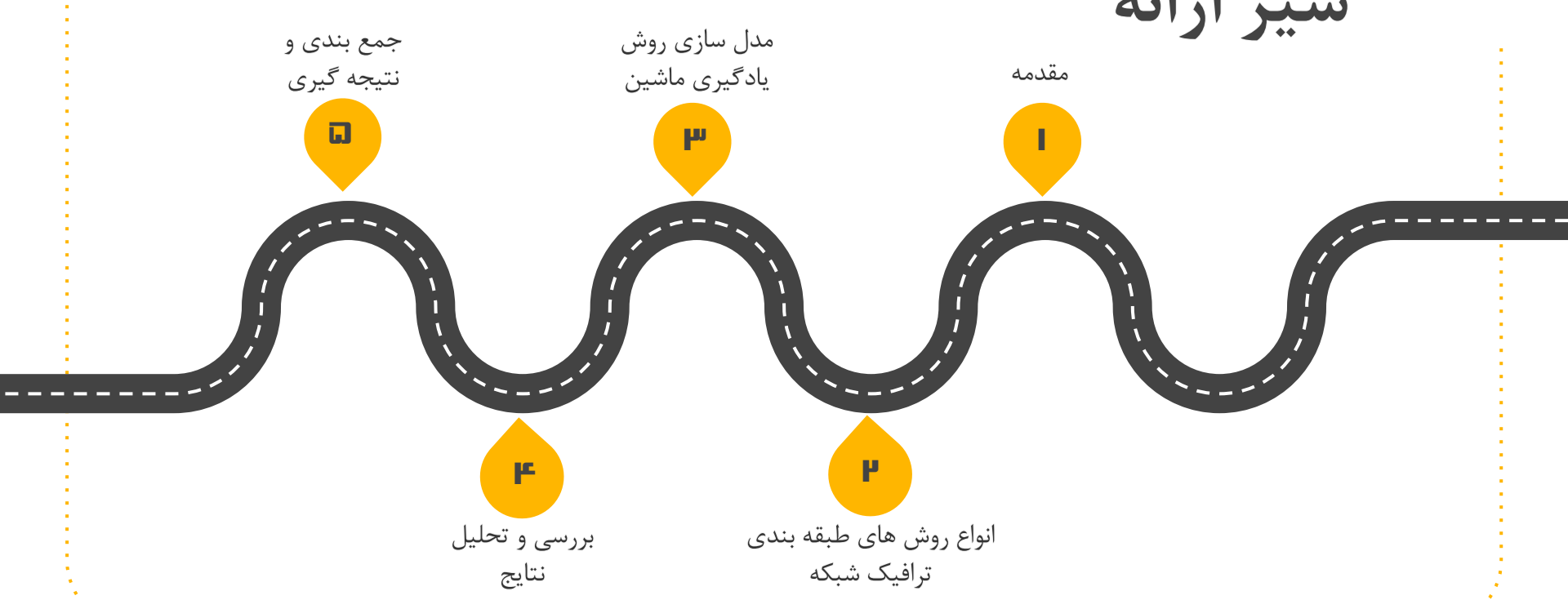
طبقه بندی ترافیک شبکه
با استفاده از الگوریتم های یادگیری ماشین

ارائه دهنده:
محمدمهدی هجرتی

استاد راهنما:
دکتر رضا صفابخش

خرداد ۱۴۰۰

سیر ارائه


مقدمه
۱
انواع روش های طبقه بندی ترافیک شبکه
۲
مدل سازی روش یادگیری ماشین
۳
بررسی و تحلیل نتایج
۴
جمع بندی و نتیجه گیری
۵

مقدمه

- افزایش روزافزون کاربران و ترافیک اینترنت
- مدیریت عملکرد کلی شبکه برای ارائه دهندگان خدمات اینترنت و اپراتورهای شبکه
- برنامه ریزی صحیح در قسمت های مختلف شبکه
  تخصیص منابع، بهبود کیفیت خدمات سرویس و ...

۲

انواع روش های طبقه بندی ترافیک شبکه

۱-۱ روش مبتنی بر درگاه

- مقایسه ی شماره ی درگاه مقصد در سرآیند لایه انتقال بسته با لیست شماره درگاه های تعیین شده در استاندارد **IANA**
- استفاده ی برنامه های نظیر به نظیر از درگاه های پویا

برنامه	شماره درگاه اختصاص داده شده
FTP Data	۲۰
FTP	۲۱
SSH	۲۲
Telnet	۲۳
SMTP	۲۵
DNS	۵۳
HTTP	۸۰
POP3	۱۱۰
NTP	۱۲۳

۲-۱ روش مبتنی پیلود

- جستجوی محتوای بسته ها برای پیدا کردن امضای برنامه های شناخته شده

P2P Protocol	String	Trans. Protocol
Edonkey 2000	0xe319010000	TCP/UDP
	0xe53f010000	
Fasttrack	“Get /.hash”	TCP
	0x2700000002980	UDP
BitTorrent	“0x13Bit”	TCP
Gnutella	“GNUT” “GIV”	TCP
Aress	“GET hash”	UDP
	“Get Shal”	

- نیازمند سیستم پردازشی قوی
- بسته های رمزگذاری شده

۳-۱ روش مبتنی بر رفتار میزبان

- الگوهای اتصال مختلف در برنامه های مختلف
- استفاده ی میزبان از یک برنامه در هر لحظه

۱-۴ روش مبتنی بر یادگیری ماشین

- آموزش یک طبقه بندی کننده یادگیری ماشین
- طبقه بندی داده های ناشناخته

- **۱-۴-۱ یادگیری نظارت شده**

پیش بینی داده های بدون برچسب از طریق داده های برچسب گذاری شده

- **۱-۴-۲ یادگیری بدون نظارت**

پیش بینی داده های بدون برچسب بدون نیاز به دانش از قبل تعیین شده

۳

مدل سازی روش یادگیری ماشین

۲-۱ جمع آوری داده

- گرفتن بسته ها
- ابزار وایرشارک و تی سی پی دامپ

۲-۲ استخراج ویژگی ها

- تعداد بسته ها
- طول هر بسته
- درگاه
- پروتکل
- ...

- ابزار نت میت و پرل اسکریپت

۲-۳ یادگیری نمونه

- نمونه گیری از داده ها
- برچسب گذاری

۴-۲ پیاده سازی الگوریتم

- با استفاده از ابزار وکا
- شش الگوریتم یادگیری ماشین:
  - درخت تصمیم آر بی اف
  - ماشین بردار پشتیبان
  - نزدیک ترین همسایه
  - سی ۴.۵
  - نیوبیز
  - شبکه بیز

۴

بررسی و تحلیل نتایج

بررسی و تحلیل نتایج

الگوریتم	Naive Bayes	SVM	Bayes Net	RBF	C4.5	NN
دقت طبقه بندی	٪۷۱.۸۹	٪۷۴.۰۵	٪۷۸.۳۲	٪۶۸.۲۵	٪۹۳.۳۳	۸۰.۲٪

جمع بندی و نتیجه گیری

- اهمیت طبقه بندی ترافیک شبکه
- روش های طبقه بندی
- مدل سازی روش یادگیری ماشین

منابع و مراجع

[1] Internet assigned numbers authority (iana). https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml. Accessed: 2021-05-01.

[2] Jamuna, A et al. Efficient flow based network traffic classification using machine learning. 2013.

[3] Shafiq, Muhammad, Yu, Xiangzhan, Laghari, Asif Ali, Yao, Lu, Karn, Nabin Kumar, and Abdessamia, Foudil. Network traffic classification techniques and comparative analysis using machine learning algorithms. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pages 2451–2455. IEEE, 2016.

[۴] خویی، زهره امینی. طبقه بندی ترافیک شبکه با استفاده از الگوریتم جنگل تصادفی بهبودیافته. ۱۳۹۶.

با تشکر از توجه شما

m.hejrati@aut.ac.ir