



网元性能及业务数据特征精准分析方法

郭宣羽

(中国移动通信有限公司研究院, 北京 100053)

摘要: 传统的网络性能分析以及业务数据分析, 都是通过网管 KPI 确定的, 网络指标出现异常则需要根据优化人员经验进行问题定位。对问题定位的效率取决于技术人员的水平, 同时网络业务数据特征的分析粒度也很粗, 不能精细化、全面化、多样化地分析用户特征。随着统一 DPI 系统的引入, 展示了用户会话过程中涉及的最细粒度的业务状态以及网络信道质量。为了更加准确地反映小区网络质量和用户真实业务体验, 所研究的网元性能及业务数据分析方法采用数据切片和关联的方式, 更加精细化地将体现网络质量的控制面信令 XDR (X data recording, 呼叫数据记录的扩展) 数据和用户面业务 XDR 数据关联在一起分析。实现分应用、分业务拓展的用户感知分析, 明确质差用户或场景等情况下 XDR 的数据特征, 从而提升网络优化效率、提高用户感知。

关键词: DPI; XDR; 数据切片; 数据关联

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019205

Accurate analysis method of network element performance and service data characteristics

GUO Xuanyu

China Mobile Research Institute, Beijing 100053, China

Abstract: The performance of telecommunication network was usually analyzed by the KPI of network management system. If there was something wrong with the KPI, it needed the network optimization personnel to position problems. The efficiency of analyzing result depended on the quality of the network optimization personnel. Meanwhile, the analysis of data performance was rough estimated. As the unified deep packet inspection could show out the smaller granularity of the service status and channel quality, the analysis method of service data and node performance was discussed. By data slice and data association of control plane XDR and user plane XDR, the performance was analyzed with smaller granularity to show the quality of the network and the experience of users. So the user experience was showed out by different applies and services with different scenes, and the data characteristics of the XDR were more clearly. Finally, it improves the efficiency of network optimizing, the user feeling and the capability of data realization.

Key words: DPI, XDR, data slice, data association

收稿日期: 2019-06-20; 修回日期: 2019-09-08

基金项目: 国家科技重大专项基金资助项目 (No.2017ZX03001028)

Foundation Item: The National Science and Technology Major Project (No.2017ZX03001028)

1 引言

为了更好地进行网络监控、优化和维护, LTE 网络引入了数据采集系统(即统一 DPI 系统)^[1], 对 LTE 网络的多个接口进行数据采集, 包括 Uu/X2/S1-MME/S1-U/S6a/S11/SGs 等接口。基于这些接口上采集的原始数据, 数据采集系统基于用户维度生成对应的用户行为记录, 称作 XDR 数据。基于所采集的控制面信令而生成的信令过程的详细记录, 称作信令 XDR; 基于所采集的用户面业务数据而生成的业务传输过程的详细记录, 称作业务 XDR^[2]。

DPI 信令数据是网络部门开展感知分析保障的应用基础, 也是中国移动各类大数据分析应用的主要数据来源。现有的基于统一 DPI 数据进行的网络监控、优化以及用户业务体验分析, 主要是单独基于信令 XDR 或业务 XDR 数据进行分析, 或者不同接口分别统计指标后联合在一起分析。

为了准确地反映小区的网络质量和用户的真实业务体验, 本文的讨论采用数据切片的方式, 更加精细化地将体现网络质量的控制面信令 XDR 数据和用户面业务 XDR 数据关联在一起分析, 实现分应用、分业务地拓展用户感知分析, 从而更加准确地呈现小区用户在网络中的真实情况。

2 数据分析方案

现有网络中, 小区级网络指标不能真实、精细地评估网络状况。同时, 测试终端难以反映商用终端在现网的真实体验。通过现网所有商用终端的统一 DPI 数据进行网络评估, 能够更精确地获得网络质量的情况。

本文对现网统一 DPI 数据的用户面 XDR 和信令面 XDR 进行切片/关联分析, 还原业务和网络状态, 定制多维度的小区级/用户级业务数据指标(包括流量特征、业务特征等)和网络数据指标(包括负载信息、用户信道信息等), 实现场景感知

以及小区级/用户级细颗粒度的精细化无线资源管理。进而生成无线网络的 QoS 保障信息, 输出实际网络中的潜在问题, 指导网络性能管理和优化; 或者输出网元性能特征和业务数据特征, 最终提升网络性能, 实现智能业务实时保障, 并确保终端用户的业务体验。

2.1 数据切片/关联方案

由于 XDR 是用户每次会话期间产生的各个接口数据, 将某用户进行一次会话时间内产生的信令面 XDR 和用户面 XDR 进行关联, 即可获得当时该用户的业务情况以及信道质量情况, 而将同一业务类型的用户或同一信道质量情况下的用户关联 XDR 进行切片统计, 即可获得一个切片时间(或切片整数倍时间)内, 该业务类型的数据特征或该信道质量情况下的数据特征。

根据不同的应用分析需求, 切片/关联方案可以有以下几种。

(1) 信令/业务流程切片方案

一个用户一段时间内, 涉及的所有接口信令面 XDR 和用户面 XDR 进行关联, 可以获得该用户的具体网络行为以及业务行为。该方法适用于用户投诉以及问题定位场景, 只涉及对 XDR 数据的关联处理, 不涉及对 XDR 数据的切片处理。

(2) 标准时间区间切片方案

将一个用户所有的用户面 S1-U XDR^[3]和 UE_MR 都按照统一的标准时间进行 5 s 切片处理, 整理出标准时间最细粒度间隔内该用户的业务和信道特征信息。经过对全网所有 5 s 级切片指标的统计后, 该方法可以给出用户、小区和网元标准时间基础上, 更细颗粒度(如 5 s 级等)的业务数据指标情况以及网络信道状况, 实现能够分应用、分业务统计业务以及分信道条件地精细化分析网元性能和业务数据特征。

(3) 业务会话级切片方案

以用户的 S1-U XDR 为基准, 将用户每一个 S1-U XDR 起止时间内该 UE 的所有 UE_MR 关



联, 给出每次会话发生时的业务状态和信道条件, 然后按照用户标识、小区标识、业务类型标识等维度进行统计分析。这种数据处理方法适用于, 业务类型更加精细化的场景, 比如按照应用类型(如业务大类、业务小类)或者应用协议(如 HTTP)分析业务数据特征和网络信道状况。

(4) 用户周期测量级切片方案

以周期性 UE_MR 为基准(通常周期设置为 5 s), 将用户面 XDR 按照 UE_MR 的起止时间进行切片处理, 得到相应 UE 测量时间内的业务情况。经过 UE 级以及小区级统计, 可以得到网络信道下的业务数据情况。这种数据处理方式适用于分析不同网络信号质差情况下的业务数据特征情况。

2.2 多维指标设计方案

基于不同的切片方案, 可以设计出多维统计指标, 具体归纳为 3 类。

(1) 一维指标, 包括覆盖、干扰、容量、质量和业务大类等。例如, 某接口统计时间内某字段分区间的采样点(占比)。

(2) 二维指标, 包括覆盖/干扰、覆盖/容量、覆盖/质量、覆盖/业务、业务/覆盖等。例如, 根据多接口关联数据生成二维指标, 不同 RSRP 区间及业务流量区间的时长、速率、XDR 数量占比等。

(3) 多维指标, 在二维指标基础上, 选择多个字段, 不同的字段 1 区间&字段 2 区间&字段 3 区间..., 求第 $N+1$ 个字段的均值/采样点数/最大值/最小值以及求和等。

2.3 分析网元标签方案

通过对任意网元下所有切片关联后的 XDR 信息进行挖掘统计, 按照网络指标门限要求进行比对, 从而给出该网元的相应标签, 以便后续进行相应的网络优化建议或者应用拓展需求。网元性能以及业务标签确定流程如图 1 所示。

(1) 采集每个用户每次业务相关的信令数据和业务数据

通信协议中定义了网络节点间通过标准的格

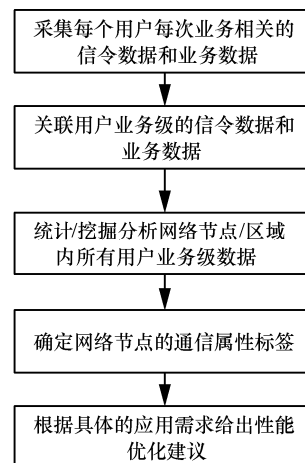


图 1 网元性能以及业务标签确定流程

式传递和解析数据, 用户在通信网络中进行业务时, 在每个接口生成相应的接口数据, 包括用户每次业务整个业务流程中产生的各接口信令数据和业务数据。获取一定时间一定范围内的所有用户每次业务相关信令和业务数据。

(2) 关联用户业务级信令数据和业务数据

通过各接口数据中的用户标识信息、网元信息、时间信息以及业务标识信息等, 将每个用户每次业务流程中涉及的相关接口数据关联作为该用户的一次业务级数据记录。

(3) 统计/挖掘分析网络节点/区域内所有用户业务级数据

根据特定网络节点或标识, 筛选出相应节点/区域范围所有用户的一定时间内的业务级数据, 通过对用户级网络指标进行统计分析, 或采用数据挖掘方法, 计算或分析得出一系列相应的统计指标数据。

(4) 确定网络节点/区域的通信属性标签

根据特定网络节点或区域内用户数据的相应统计指标结果以及挖掘结果, 结合通信各个特性的相关门限要求, 对比并确定网络节点/区域的多个通信属性标签。

(5) 根据具体的应用需求给出性能优化建议

根据业务需求以及通信知识, 综合分析网络节点的多个通信属性标签, 最终确定该应用需求的优化建议方案。

3 数据特征分析应用举例与验证

3.1 低速率小区数据特征分析

以分析低速率小区为例验证基于 S1-U XDR (包括所有业务类型) 级别的下行流量以及业务时间, 统计小区级的平均下载速率。对筛选出的低速率小区 XDR 进行分析, 根据经验门限为小区打标签, 给出可能导致低速率问题的原因。

例如, 数据特征分析过程可如下。

(1) 筛出低速率小区, 基于 S1-U XDR 业务类型数据表单, 统计所有 XDR 的下行流量以及业务时长, 统计所有小区的平均下载速率, 并筛选出平均下载速率小于低速率判决门限的小区。

(2) 关联 UE_MR 与 S1-U XDR, 基于 cellID、IMSI、开始和结束时间将筛选出的小区对应的 UE_MR 与 S1-U XDR 数据进行关联, 生成该 UE 的一条 S1-U XDR 时间内业务的综合 MR 信号质量情况, 或者该 UE 一条周期性 MR 时间内的信号质量对应的业务综合传输情况。

(3) 对关联后的下行流量、时长、RSRP、业务类型等字段数据进行分布统计, 给出占比超过门限的场景。

(4) 下行流量>门限及 RSRP>门限的时长占比低于门限, 则为该低速率小区打上“强覆盖区域高流量低时长”的标签。

(5) HTTP 业务响应时长>门限的比例高于门限, 则该小区的数据特征包括“HTTP 业务响应时间过长”。

(6) 分业务类型进行速率统计, 若低速率业

务流量/时长的比例高于门限, 则该小区的数据特征包括“低速率业务占比偏高”等。低速率小区高时延业务数据分析见表 1。

3.2 业务类型分布情况分析

基于给定小区 S1-U XDR 中所有业务类型的 XDR 进行小区级分析, 给出各种业务的占比情况, 根据相应的 S1-U XDR 和 UE_MR 关联数据。对小区的每一类业务分布进行分析, 给出小区的业务属性标签以及业务分布区域等标签。图 2 中, 该小区购物类占比最大, 为 28.44%, 其次是微博为 17.89%、即时通信为 15.63%和浏览下载为 14.88%。浏览下载业务大类中, 进入头条占比最大为 22.91%。这样就可给出该小区的应用大类和小类数据特征标签, 用于后续的应用分析。

另外, 也可对区域内多个小区的业务大类时长和流量占比进行对比分析, 如图 3 所示。每个小区的各种业务时长占比以及流量占比情况, 明显网购业务的占比最大。

4 结束语

本文通过对统一 DPI 系统架构和 XDR 数据格式及特征的研究, 提出数据切片关联的方案, 将每个用户的多接口业务数据和信道条件数据结合, 设计出覆盖、干扰、质量、容量和业务方面的网元级多维指标, 给出多维度不同区间内数据特征, 精准分析出用户、网元的性能以及业务特征, 从而实现能够分应用、分业务拓展的用户感知分析, 明确质差用户或场景等情况下 XDR 的数据特征。最终, 提升了网络优化效率、提高了用

表 1 低速率小区高时延业务数据分析

小区标识	下行流量区间/KB	XDR 数	UE 数	XDR 数占比	平均业务响应时长/ms
100668929	(~,500]	6 490	427	13.14%	76.4
100873473	(~,500]	5 495	324	17.57%	114.8
101018881	(~,500]	3 812	172	15.09%	224.6
101246209	(~,500]	2 210	204	13.73%	108.3
101350924	(~,500]	1 079	177	10.09%	57.4

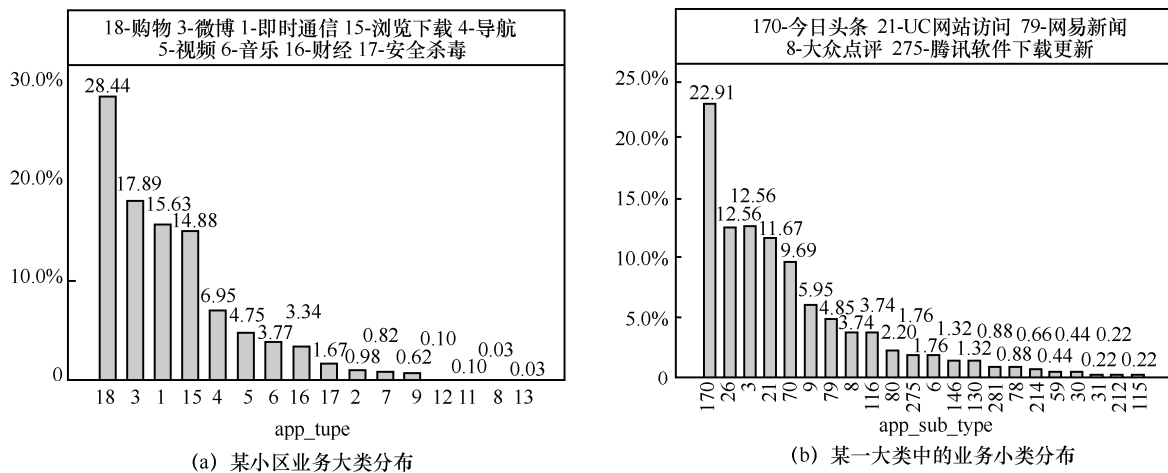


图2 某小区业务大类分布以及某一大类中的业务小类分布

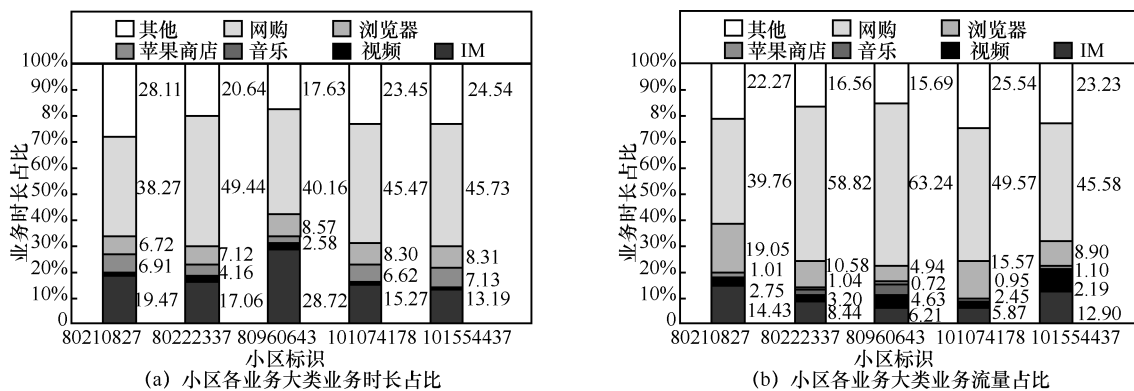


图3 小区各业务大类业务时长及业务流量占比

户感知以及增强了数据变现能力^[4]。

本文的系统平台已接入现网数据进行试点验证,证明了该体系的可实施性和可部署性,可保证3种切片/关联数据方案的准确率达到100%,正确实现性能分析方案,给出异常问题小区以及异常点。根据引入的数据源不同,本系统同样适用于5G网络性能评估、NB-IoT网络性能评估、基于用户感知的用户级分析等。

参考文献:

- [1] 中国移动. 中国移动统一 DPI 设备技术规范-总体技术要求规范 v2.4.0[S]. 2018.
China Mobile. China Mobile unified DPI equipment technical specification - general technical requirements specification v2.4.0[S]. 2018.
- [2] 中国移动. 中国移动统一 DPI 设备技术规范-信令采集解析服务器接口规范 v2.4.0[S]. 2018.
China Mobile. China Mobile unified DPI equipment technical

specification-signal acquisition and analysis server interface specification v2.4.0.[S]. 2018.

- [3] 中国移动. 中国移动统一 DPI 设备技术规范-互联网与移动互联网数据面部分[S]. 2018.
China Mobile. China Mobile unified DPI equipment technical specification - data plane of internet and mobile internet [S]. 2018.
- [4] 谷红勋, 杨珂. 基于大数据的移动用户行为分析系统与应用案例[J]. 电信科学, 2016, 32(3):139-146.
GU H X, YANG K. Mobile user behavior analysis system and applications based on big data[J]. Telecommunications Science, 2016, 32(3): 139-146.

[作者简介]



郭宣羽(1982-),女,中国移动通信有限公司研究院工程师,主要研究方向为4G/5G无线网络技术、统一DPI系统研究、无线网络性能评估等。