

硕 士 研 究 生 读 书 报 告



题 目 基于MQTT的移动应用消息推送

作者姓名 阮野

作者学号 21651044

指导教师 李启雷

学科专业 移动互联网与游戏开发

所在学院 软件学院

提交日期 二〇一七 年 一 月

Mobile Application’s Push Message Based on MQTT

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Qilei Li

By

Ye Ruan

Zhejiang University, P.R. China

2017

摘 要

随着智能设备的快速普及和移动应用的迅猛发展，已进入移动互联网时代。以 iPad、iPhone 以及 Android 手机为代表的移动设备和智能终端近年来取得了巨大的发展，其用户数迎来了爆发式的增长。为了适应实际应用中对信息获取及时性，耗电量，以及网络环境等方面的需求，需要用推送的方式取代传统拉取的方式来进行消息的传递。MQTT(Message Queuing Telemetry Transport, 消息队列遥测传输)就是专为这种情况所设计的一种消息传递协议。本文介绍了MQTT 协议的基本内容和特点。文章讨论了使用MQTT协议构建消息推送服务的必要性和适用性，并指出了在具体实现上应注意的一些关键问题。

**关键词**：**MQTT协议，消息推送，移动应用**

Abstract

With the rapid popularization of intelligent devices and the rapid development of mobile applications, we have entered the mobile Internet era. Mobile devices and smart terminal represented by iPad, iPhone and the Android phones have made a significant development in recent years. The number of their users is growing explosively. More and more people, especially the younger generation use the intelligent Terminal instead of the traditional PC terminal to obtain information. In order to meet the requirements of timeliness of accesses of information, power consumption, as well as the network environment in practical application, we should replace the traditional pull way with the push way to deliver messages. MQTT (Message Queuing Telemetry Transport) is a messaging protocol just designed for the situation. This paper describes the basic content and features of the MQTT protocol. This paper discusses the necessity and applicability of using MQTT protocol to construct message push service, and points out some key problems that should be paid attention to in the concrete realization.

**Keywords**：MQTT, Message pushing, Mobile Applications

## 一．背景

根据中国互联网络信息中心发布的《第 32 次中国互联网络发展状况统计报告》显示，截至 2013 年 6 月底，我国网民规模达到 5.91 亿，我国手机网民规模达到4.64 亿，网民中使用手机上网的人群占比提升至 78.5%，移动互联网用户已成为推动中国互联网深入普及的关键动力，人们的生活服务开始移动互联网化，从而使得移动互联网应用领域进一步拓展，从娱乐主导型向消费和电子商务型转变。随着移动互联网和智能终端的普及，移动社交网络，如微信、新浪微博、人人网、Facebook 等迎来了爆炸式的发展，这对移动终端设备上信息的获取提出了较高要求，主要体现在移动性和时限性两个方面。移动性，要求低功耗、低速率条件下的消息传输; 时限性，要求信息在规定时间内发送到移动终端设备上。

目前，用户在移动终端上获取信息主要有“客户端拉取”和“服务器推送”两种方式。在拉取的方式下，用户首先确定自己的信息需求，再通过客户端向服务器发送请求，进行有目的的查询和搜索，服务器只是被动的接收客户端的查询请求并经行响应。以手机上访问网页所使用的 HTTP 协议为例，当用户通过Internet 访问某一网址的服务器的时候，客户端与服务器先建立好连接，然后根据用户的输入产生 HTTP 请求，发送给服务器。服务器端接收到请求后，产生相应的响应数据，并封装成 HTTP 格式的数据包，返回给客户端。用户的浏览器接收并解析返回的数据后，将结果显示给用户，并且关闭当前的连接。拉取的主要优点是：针对性好。用户可以针对自己的需求有目的地去查询、搜索用户所需要的各种信息。但是这种方式的缺点也很明显：实时性差。当系统中与用户有关的信息发生变动或更新的时候，用户无法及时的拉取最新的动态信息。为了提高及时性，用户必须尽可能的缩短发送查询请求的时间间隔，这样就加重了客户端的负担，并且造成大量的网络流量的浪费，在网络不稳定的移动终端上，这样的弊端尤其明显。

服务器推送技术是指服务器将用户所需的信息主动推送给用户而无需用户自己进行拉取操作。这种方式的优点是：

1．及时性好。因为信息是由服务器主动推送的，所以客户端可以在第一时间收到服务器发来的更新的消息或者是更新提醒，更能满足用户的需要。

2．效率高。相比拉取技术普遍采用短连接来说，推送技术一般采用长连接的方式，客户端在与服务器的通信期间一直保持这个长连接，而无需耗时间为后续数据建立新的连接，同时由于客户端不用发送查询请求，进一步减少了网络中的数据流量。

由于具有以上的优点，在移动互联网中，消息推送技术必将成为未来发展和应用的主要技术。推送技术虽然在客户端上的效率更高，但它增加了服务器端的负担，对服务器的性能和效率提出了更高的要求。同时，系统还要面对不断增长的用户数量和各类信息和数据的处理，经过一段时间的运行后，往往会出现服务器负载过重，对用户请求的响应速度过慢的情况。传统的处理方法中，一般都是对服务器进行软件或者硬件上面的升级，比如更换服务器的操作系统，增加服务器的容量，替换成处理速度更快的服务器等。但这样做会带来两个方面的问题：一是在软硬件的升级过程中，整个系统的服务必须停止，这对于某些不可以中断的业务而言将造成很大损失，同时也会影响用户的体验和使用。二是在升级硬件之后，原有的硬件将被闲置，造成资源上的浪费。

## 二．发展现状

推方式用到的推送技术（Push Technology）是由美国 PointCaseNetwork 公司在 1996 年提出的，目的在于提高信息通过计算机网络的获取效率。随着移动互联网在人们移动办公、学习、娱乐等方面的深刻影响，推送技术的运用也愈发广泛，国内外业界对推送技术都有大量研究与实践应用。

移动推送技术是由国外兴起并发展起来的，因此运用该技术的内容和形式也是多种多样，并且在新兴的物联网、云服务等领域发挥越来越重要的作用。在推送协议的规范制定方面，经过多年的积累和发展，形成了主流的 XMPP 协议和逐渐成熟的 MQTT 协议两大类，这两类协议介绍如下：

（1）XMPP：XMPP（可扩展消息处理现场协议）是基于可扩展标记语言（XML）的协议（类似于 HTTP 协议的数据传输协议），前身是 Jabber，最早版本由 Jeremie Miller 在 1998 年发布，目前已由 IETF 国际标准化组织完成了标准化工作，正式列为认可的实时通信及 Presence 技术，XMPP 的技术规格已被定义在RFC 3920 及RFC 3921 中。优点是协议成熟、强大、可扩展性强、目前主要应用于许多及时通讯软件中，缺点是协议较复杂、冗余（基于XML）、费流量、费电，部署硬件成本高。

（2）MQTT：MQTT（Message Queuing Telemetry Transport，消息队列遥测传输）是IBM开发的一个即时通讯协议，有可能成为物联网的重要组成部分。该协议支持所有平台，几乎可以把所有联网物品和外部连接起来，被用来当做传感器和致动器（比如通过Twitter让房屋联网）的通信协议。MQTT 协议是一种轻量级基于代理的发布/订阅的消息传输协议，其设计思想是开放、简单、轻量、易于实现，适合在带宽、计算和处理能力受限的环境下工作。这种特点使得 MQTT 协议非常适合在移动应用中使用，并且相比其他移动通信协议，MQTT 协议有着较大的优势。MQTT 是开源的协议，当前协议版本为 V3.1，由 IANA 保留的标准端口号为1883和8883（支持 SSL），2014 年成为OASIS推荐的物联网传输协议标准。

在提供推送服务的应用框架方面，有 Google 推出的 GCM（Google Cloud Messaging）服务，它是一个能够帮助开发者从服务器端发送数据到运行在 Android手机上的程序的服务，可以处理所有的消息队列的问题并且可以把消息发送到目标机器上运行的目标程序，其中采用 XMPP 协议来保持 Android 设备与云服务器的连接，该服务优点是完全由 Google 提供服务，完全免费，无需实现和部署服务端，缺点是该服务在国内不够稳定，需要用户绑定 Google 帐号，受限于 Google的规范；与 GCM 类似的推送服务还有 APNS（Apple Push Notification Service）服务，它是 Apple 公司在 2009 年发布 iOS3.0 中的一个新功能，通过终端设备与推送服务器建立持续的 IP 连接，将来自第三方应用服务器的提示信息推送到iPhone、iPad 和 iPod Touch 等终端上，其使用的协议也是基于 XMPP 协议；此外，市场上存在有丰富的第三方推送服务商，例如有较著名的 Urban Airship，它提供实时丰富的推送通知，允许在通知中包含 html、音频、视频、地图以及其他富媒体内容，实时跟踪推送通知的效果，该服务可在 Android、iOS、Windows8 和Blackberry 上运行。

在企业级推送解决方案方面，有 IBM 的 Worklight，它是面向智能手机和平板电脑的开放性移动解决方案，是一个可以利用第三方库和框架不断发展的生态系统，能连接各种后端企业系统和基于云的服务，并可以从一个中央界面控制并管理各种应用程序组合，可为各类企业提供的跨平台解决方案。此外，还有 Sybase公司的 SUP（Sybase Unwired Platform），它是企业级移动应用平台，支持企业实现应用程序移动化的体系架构，帮助企业将适当数据和业务流程移动化到任何移动设备上，提供从信息中心向包括员工和最终用户在内的移动终端发送信息的功能，协议是基于 HTTP 或 HTTPS 方式完成传输，适用于大中型企业的复杂业务解决方案。

关于信息推送的研究国内起步较晚，根据相关资料显示国内是从 2001 年逐步增多对此技术和应用的研究，研究范围主要集中在运用推送技术，随着近年来国内移动互联网的迅猛发展，该技术在教育、医疗、媒体、电商等行业的运用也越来越多。

在提供推送服务的应用框架方面，有阿里巴巴支持开发的 RocketMQ 开源项目，它是一款分布式、队列模型的消息中间件，具有亿级消息堆积能力和比较友好的分布式特性，同时支持 push 与 pull 方式消费消息，用于阿里巴巴消息服务产品以及作为第三方软件产品的消息中间件。

在企业级推送服务解决方案方面，国内也有众多第三方推送服务商，其中的个推是国内最大的推送技术提供商，以 SAAS 模式为企业提供长连接 SDK 和服务器接入的整体解决方案，实现对用户的免费实时推送通讯服务，支持iOS和Android系统。类似的第三方推送服务商还有极光推、华为 Push、百度云推送等。此外，还有以腾讯公司的微信为代表的消息和通讯服务平台，微信是于 2011 年推出的即时通讯服务的软件，公众账号可向关注者提供信息推送的服务，所使用的协议为结合 XMPP 和类 Sync 的自定义协议。

目前国内针对中小型企业的解决方案还处在发展状态，而且主要集中在互联网行业客户，由于中小型企业需求相差较大，很难有统一认可的较完善的解决，因此市场虽然有众多的第三方推送平台却很难适合各类行业的需要。

## 三．MQTT协议

消息队列遥测传输( MQTT, Message Queuing Telemetry Transport)协议是IBM公司提出的开放协议，最初的设想是应用于大量计算能力有限的传感器等微型设备，其工作的网络带宽低且不稳定，但又需要保证网络节点之间的可靠通讯。

该协议目前已经被结构化信息标准促进组织(OASIS)接受，并将其建议为物联网消息传递协议的首选标准。MQTT已经成为物联网领域的事实标准，IBM 公司已经成功将其应用于智能实验室、远程医疗中心等项目，并推出了MessageSight等MQTT中间件产品，其它企业和机构也相继跟进，发布支持MQTT协议的开源/商业产品，或采用MQTT协议构建相关应用。伴随着物联网的迅猛发展，MQTT协议未来的发展不可限量。

由于物联网传感器等设备的计算能力和电量都非常有限，因此MQTT的设计理念从一开始就是简单，轻量，节省电力，这正好满足了移动应用的一个重要诉求。当前电池储能技术的发展远远不能满足智能手机的发展需要，使用智能手机用户最痛苦的事情就是手机耗电太快，手机的电量甚至不能支持完8小时的工作时间。但如果要保证消息及时到达，手机就必须时刻保持与服务器的连接，并定期与其通讯，检查是否有新消息。MQTT协议非常精简，额外的数据传输量非常小，可以在最大限度节省电力，同时也节省网络流量。这是选择MQTT构建消息推送服务的最大原因。

3.1 MQTT内容

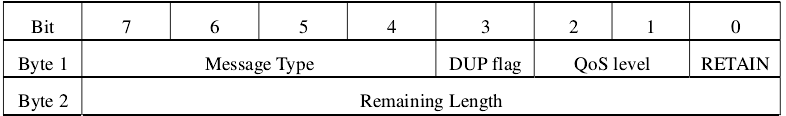
每一个 MQTT 命令消息的消息头里都包含一个固定报头，表3-1 展示出固定报头的格式。

表3-1

如表中所示，固定报头分为2个字节，第一个字节包含消息类型和标志字段（DUP、QoS level、RETAIN），第二个字节包含剩余长度字段（至少一个字节的长度）。其中的 Message Type 占第一个字节的4到7的比特位，表示为一个4比特位的无符号值，协议中消息类型定义如下表 3-2 所列。第一个字节剩余的比特位包括表示 DUP、QoS 和 RETAIN 的字段，DUP 在第 3 个比特位上，当客户端或服务端试图重新传送 PUBLISH、PUBREL、SUBSCRIBE 或 UNSUBSCRIBE 命令消息时设置此标志。这适用于消息的 QoS 值大于 0，以及需要确认的情况。当 DUP 位被设置时，可变报头应包括一个消息 ID。接受者应把这个标志作为是否先前已经收到了消息的提示，但不应该用来检测重复情况。QoS 位于第一个字节的1和2比特位上，这个标志表示传递一个 PUBLISH 消息所确保的一个级别。RETAIN在第一字节的0比特位上，这个标志只用在 PUBLISH 消息上，当客户端发送一个PUBLISH 到服务端时，如果该标志位置为 1，服务器应该保持该消息信息直到它已传递到当前订阅的用户那里。当该标志位置为 1 并且一个主题有了新的订阅者，那么该主题最近的保留消息就会发送到该订阅者那里去，如果没有保留消息则不发送。Remaining Length 剩余长度，在第二个字节上，表示当前消息剩余的所有字节数，包括可变报头和负载的数据。该标志可扩充到最多 4 个字节，允许一次消息发送的最大字节数为 268 435 455 (256 MB)。

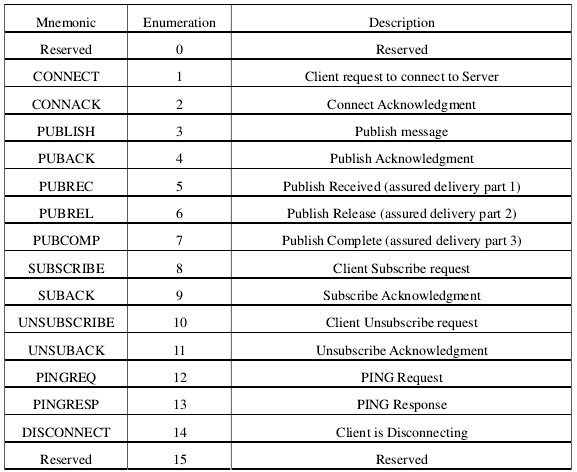


表3-2

一些类型的 MQTT 命令消息还包含一个可变报头组件。它位于固定的报头和

有效载荷之间。可变字段按顺序包括如下内容：

（1）Protocol name 该字段出现在 CONNECT 消息的可变报头中，内容为UTF编码的字符串 MQIsdp。

（2）Protocol version 该字段出现在 CONNECT 消息的可变报头中，用 8 比特的无符号值表示客户端使用的 MQTT 协议版本号。

（3）Connect flags 总共一个字节，包括 Clean session flag 占其中一个比特位，用来设置是否保存客户端的连接信息以及 QoS1 和 QoS2 等级的消息，方便重新连接后继续使用；Will flag 占其中的 2 个比特位，遗嘱消息定义为如下情况，一个消息正从服务器发布到客户端时，发生了 I/ O 错误或客户端无法在心跳时间内通信；Will QoS 占其中 4 和 3 比特位，用来设置 Will 消息的发送等级，分为 0、1、2 三个等级；Will Retain flag 占其中第 5 比特位，用来设置服务器是否保存 Will 消息；User name and password flags 占其中的 7 和 6 比特位，用来设置在 CONNECT 时是否使用用户名和密码。

（4）Keep Alive timer 心跳时间，在发送 CONNECT 消息时设置，客户端与服务端需在心跳时间内完成 PINGREQ 和 PINGRESP 的交互确认，如果在一个半的心跳时间后还无法完成交互确认，服务端会与客户端断开连接，其占用 2 个字节，单位是秒钟，最大值约为 18 小时。

（5）Connect return code连接返回码，用在CONNAK消息中，占用一个字节，0表示成功，其它表示各类连接状态。

（6）Topic name 主题名称，用在 PUBLISH 消息中，是信息频道的标识，是发布信息，订阅频道的关键基础，主题名称是用来标识该发布哪个信息频道的有效载荷数据的关键。订阅者可标识他们想要接收的消息的信息频道。主题名称的最大长度限制为 32767 个字符数。

除了短小精悍，节约带宽，节省电量消耗，MQTT协议还具有其它的特性，非常适合用于构建移动应用的消息推送服务。

（1）保证服务质量。MQTT提供三种消息传输的服务质量保障水平，用户可以根据实际需要进行选择：

QoS0：至多传递一次。消息有可能会丢失。这种服务质量下，消息传递的速度最快，但不能保证消息的可靠到达。通常适用于网络环境差，而且不在意单次数据丢失的情况，例如GPS数据的采集。

QoS1：至少传递一次。可以确保消息被可靠传递到目标，但可能会有消息被重复传递。这种服务质量权衡了传输效率和消息可靠性，最常被采用。

QoS2：确定只传递一次。消息不会被丢失，也不会被重复传递。这种服务质量被用来保证最高的消息传递服务质量。

（2）连接中断通知。与物联网类似，手机的网络质量远不能与连接网线的台式机相比，当我们发生位移时，或者进入电梯，地下室等区域时，手机经常会发生基站切换，网络信号消失，连接中断等情况。MQTT协议可以很好地适应这种不稳定的网络环境，并且保证数据的可靠传输。而且，当发生网络异常中断时，MQTT还支持遗嘱机制，将网络中断事件通知给指定的目标对象。

（3）信息广播机制。MQTT采用发布/订阅机制来传递信息，多个手机终端可以订阅同一个主题；服务端只需要针对主题发布信息，所有订阅者都可以收到，而无需对逐个手机发送，简单高效。例如，股票价格信息，不同的人可以订阅关注同一个股票的价格信息，当该股票的价格发生变化时，服务端只需发布一次最新价格，所有的订阅者都会收到同样消息。

其他问题

在采用MQTT协议构建移动应用的消息推送服务时，也应当了解MQTT的不足，并妥善处理如下问题。

避免客户端越权订阅：

MQTT采用发布/订阅模式处理消息传递，消息是针对主题传递，而不是目标客户端。客户端只要进行订阅，就可以收到消息。理论上，如果某个客户端订阅了根主题，就可以收到所有人的消息，这是不可接受的。可以启用MQTT的用户认证机制，只有经过认证的用户才可以订阅主题，但这只能防止非认证用户的恶意订阅，对于认证通过的用户则不起作用。彻底的解决方法需要在服务端进行控制，不允许用户订阅其没有权限的主题。另外，一些MQTT产品对此也有处理，例如，IBM的MessageSight产品就支持点对点和发布/订阅两种方式，而使用点对点方式时，任意订阅都不起作用。

可扩展的MQTT服务端：

消息推送需要移动设备与推送服务器保持长连接，这与传统应用不同。传统的B/S应用是短连接模式，客户端向服务端发起请求，服务端处理完成后返回结果给客户端，然后释放连接。如果不考虑业务逻辑处理对服务器资源的使用，单个服务器可以轻松支持几十万，甚至上百万的用户访问。但是，如果拿一台普通的服务器当做消息推送服务器，单台服务器的连接能力大约在2万到10万之间，如果要保证稳定的服务能力，客户端的连接数通常不允许超过5万。IBM的MessageSight可以支持100万的客户端连接。但不论其数值多少，单台服务器的处理能力总有极限。如果要推送的设备数量超过了单台的处理能力，就需要考虑服务端集群。MQTT在服务端的集群扩展方面没有规定，需要自行设计实现水平扩展架构。

参考文献

[1]蒋鹏,袁嵩. 基于MQTT协议的综合消息推送[J]. 现代计算机(专业版),2014,11:11-15+21.

[2]关庆余. 基于MQTT协议的移动社交网络客户端的设计与实现[D].中国科学院研究生院（沈阳计算技术研究所）,2014.

[3]朱艳. 移动应用的消息推送与MQTT协议[J]. 无线互联科技,2015,08:1-3.

[4]彭喜燕. 基于MQTT的智能化物流信息系统[J]. 电脑与电信,2016,Z1:75-77.

[5]杨鹏. 基于MQTT协议的信息推送平台系统的设计与实现[D].电子科技大学,2015.