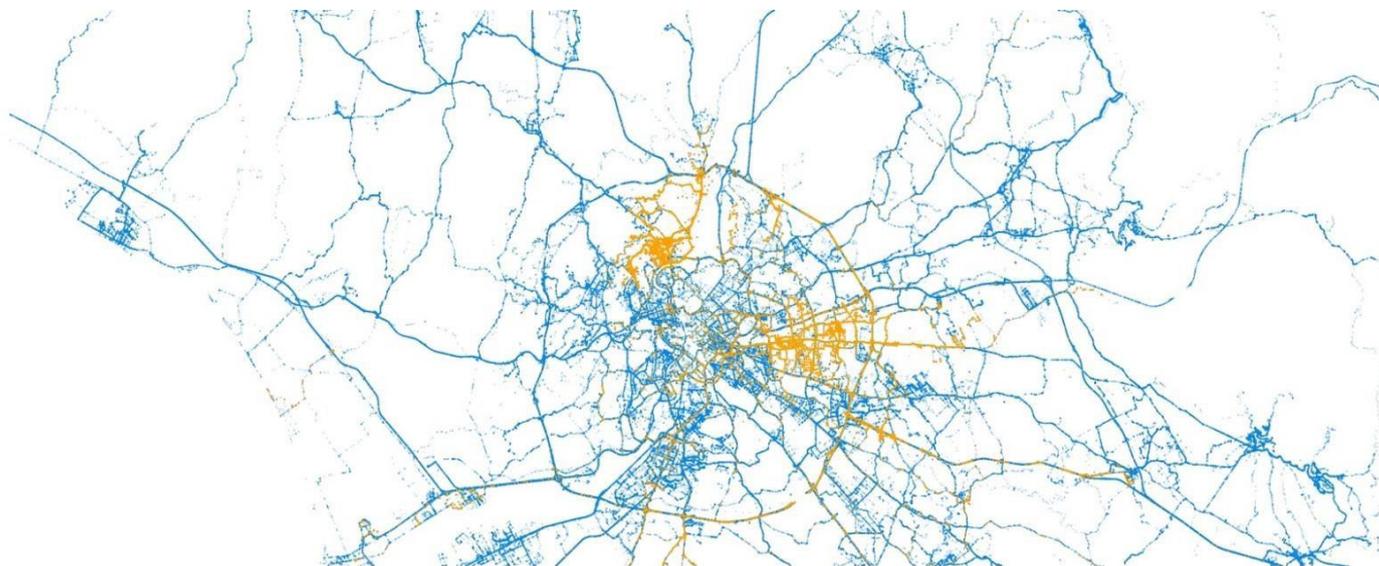


欧洲保险业领导者与明德合作分析车联网数据

并检测驾驶行为

Remi BELLINA, IA
Antoine LY, IA
Fabrice TAILLIEU, IA



近年来数字革命颠覆了各行各业，并且 5G 技术将会加速这一趋势。公司需要调整业务模式以充分考虑技术进步的影响并提高生产效率。在用户越来越期待信任、响应速度和透明度的今天，通过服务创新满足客户新需求往往是真正的挑战。保险业并不能免受这场数字革命的影响，而车联网则是保险科技公司发起的这场技术变革的极好例证。

汽车保险通常被称为“消费者诉求产品”，在这一竞争激烈的市场中，利润边际管理至关重要。本文试图重点讨论汽车保险和车联网的未来，并就明德分析团队领导的一些项目做一个介绍。

欧洲车联网：由意大利主导的异质市场

除保险之外，驾驶习惯和道路使用情况也持续发生变化：职业司机提供的车辆服务、拼车、电动汽车、新的城市交通方式、无人驾驶等。尽管车辆安全性不断提高，并且监管机关正在研究速度管制，车险理赔成本仍居高不下。外部形势持续变化环境中，更好的理解风险因素并控制成本对保险公司来说至关重要，车联网则可以提供切实有效的解决方案。越来越多的国家正推出新的数字化保险产品，例如按里程收费（PAYD）：里程越少的驾驶员保险费就越低；或者按驾驶行为收费（PHYD）：奖励驾驶技能好的驾驶员。虽然有几个欧洲国家在使用这些产品，但这些产品并不同质，各个市场发展情况也各不相同。

图 1：市场研究的随机样本



意大利拥有超过 500 万用户，是最成熟的车联网市场之一，和英国一起居于欧洲市场领先地位。他们的邻国似乎正在效仿这一作法，但里程保险（UBI）普及率仍然较低。在欧洲部署车联网系统并非一帆风顺，主要是因为设备成本挤压狭小的利润空间；在价格竞争激烈的市场拓展也较困难；在数据保密敏感的市场投保人和保险公司的之间没有足够的信任（欧洲市场的通用数据保护条例[GDPR]）。投保人不希望被监视或者因为不恰当的原因导致保险费率增加，或者索赔被拒。

产品设计和市场营销的思路也因不同市场的文化差异而显著不同。例如，按里程收费在英国很受年轻驾驶员的欢迎；按驾驶行为收费的产品在意大利则更畅销；而在瑞士，车联网则更多地用于安保，用于针对豪华轿车的增值服务；在法国、德国和西班牙，用户则对他们认为具有侵入性的解决方案不感兴趣，目前在这三个国家，包括专业车队在内，也只装备了数万辆汽车。

意大利是迄今为止欧洲最大的车联网市场，最初出于减少欺诈事件而安装车联网设备，现在有更广泛的应用，比如提升的风险认识和新服务设计方面都确实很有前景。车联网发展的趋势正在发生变化，欧洲的车联网项目正在增加，并且在加速推进。

更多的技术选择和更广阔的机会

智能手机应用是最方便部署的方式。利用智能手机的集成组件，软件开发工具包 (SDK)，可以快速部署车联网，同时还能抓取驾驶员的分心驾驶情况¹。但是，移动设备的多样性 (GPS 和加速度计不尽相同) 和用户使用情况的不确定性 (用户可能会选择性停用) 有时会导致车联网服务提供商倾向使用与汽车直接连接的组件，如加密狗(dongle)和内置盒子。通常用户可以直接将加密狗连接到汽车上，内置盒子则需要专业人员进行安装。这两种方式都更难断开连接，也能够收集完整的驾驶信息，因此对提供车辆救援服务非常有用，但它们初始的固定部署成本更高，尽管近来部署成本有下降的趋势。市场上有些解决方案融合智能手机应用和内置盒子。

无论选择哪种设备，都可以收集驾驶员行为数据 (DBD)。这些数据可以提供“时间、地点和方式”等有用信息，帮助更好地分析驾驶相关的风险。

具体来说，车联网设备可以提供大量的新变量行驶距离、加速度、转弯，地理位置等)，以便保险公司用于驾驶评分、定价、核保、理赔管理或车辆救援。

图 2：车载信息分析简图



¹ Scott, S.L. (2017 年 8 月 22 日)，Zendrive 采集 300 亿英里的智能手机数据，并与明德合作构建业内顶级的预测模型。

明德白皮书于 2018 年 5 月上传 <http://www.milliman.com/driving-risk-score/>。

根据以往的分析，按里程收费和按驾驶行为收费的保险产品是最成熟的车联网保险产品，使得驾驶员支付和个人风险特征更匹配的保费成为可能。然而我们发现新的模式主要侧重于改进服务以及提升与投保人的关系：包括个性化救援服务、父母对年轻驾驶员的控制、更透明的驾驶概况、车辆的预测性维护，简化和快速的理赔处理等。最终甚至可应用于检测特定路段的异常指标建议路况维护。车联网带来无限可能，关于车联网的讨论也逐渐清晰：为投保人提供服务和预防措施的重要性绝对不低于个性化的保险定价。

车联网是车辆保险的未来，远远不止定价，而且可以预防投保人风险。控制和降低风险是保险的本质。这是保险公司重返核心业务，回归“保险姓保”的重要机会。

抽取并强化相关信息

对于许多保险公司而言，车联网数据意味着新的原始数据材料。每天检索的数据包括：带时间戳的 GPS 坐标、加速度计和陀螺仪。需要理解和分析这些新的数据源，以便设计 UBI 里程保险车联网评分等相关指标。

必须使用统计方法处理大量数据以推断总体的平均趋势，此外也必须监测和理解某些路程的特殊性，以确保做出正确的解释。数据的认识 and 解读是营造优质客户体验的基石，因此需要综合各种算法来探索和分析数据。

传统的保险数据是“静态数据”，主要侧重于车辆和从人相关信息，车联网数据则直接处理车辆移动和驾驶行为。对保险公司来说，这是一个可能让保险公司熟悉实际路况的机会，获取的变量至少在直观上更有助于解释理赔经验。这之间可能没有因果关系，但有实证的相关性，让用户更容易感觉到可以为自己的风险行为“负责”：传统保险中年龄常被用来决定保费，客户无法改变其年龄，但是客户可以提高自身的驾驶行为，这样客户不只是单纯付费，也可以更好的理解他们支付的保费。在保险科技热潮中，客户处于商业模式的核心地位；保险公司当然也可以发挥重要作用，也有更多机会与投保人多方面接触，而在传统的业务模型里面，保险公司和客户的接触仅限于客户索赔。

掌握原始信息数据

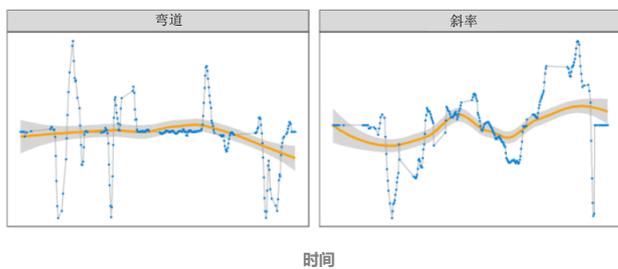
保险公司、车辆制造商和其他供应商必须控制车辆网数据，以便设计基于车辆使用的服务。

图 3: 数据示例

纬度	经度	海拔	时间
度	度	米	日/月/年 时:分:秒
48.8755	2.2848	50	21/05/2018 20:30:00
48.8760	2.2868	55	21/05/2018 20:30:05
48.8757	2.2884	60	21/05/2018 20:30:10
48.8755	2.2905	65	21/05/2018 20:30:15
48.8753	2.2919	75	21/05/2018 20:30:20

从带时间戳的 GPS 坐标中可以提取多个指标：平均速度、瞬时加速度近似值、弯道半径、里程等。下图显示经计算和平滑处理后某个特定轨迹随时间变化的相关数值。

图 4: 随时间变化的路况指标

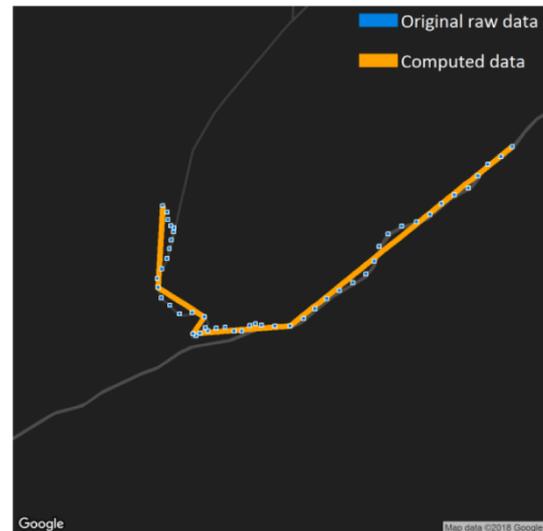


此外，还可以获得其他数据，如人口密度、气象条件，甚至该地区的一般地形。综合所有这些特点，就可以获得驾驶员或路程的概况。此外，加速度计和陀螺仪可以作为这些数据的补充，特别是用于事件分析的时候。

Python 或 R 等低成本开源解决方案通常用于处理这些变量，以便理解和研究所收集数据。面对大规模数据时，也会使用 Hadoop/Spark 等大型数据基础设施和传统的机器学习算法提供解决方案（监督或非监督学习）。

如下图所示，数据再处理通常涉及 (a) 归纳和 (b) 强化。如下使用谷歌的程序接口，希望能借助归纳取得综合高速公路驾驶情况，或借助强化精确探测特定不规则道路情况。

图 5: 数据处理技术示意图



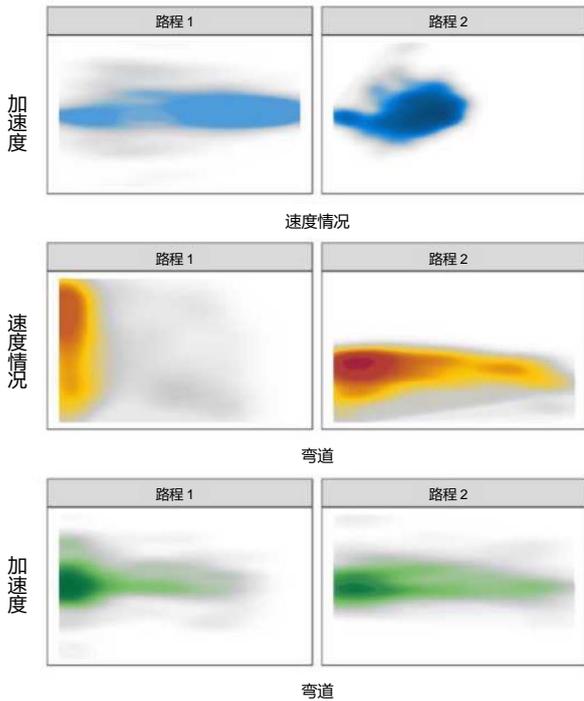
(a) 归纳



(b) 强化

数据再处理后即进行数据分析。明德的数据科学家利用数理方法，帮助保险公司研究驾驶行为：特定驾驶员的“行驶轨迹”如何？下图利用初步研究得出的适应性指标，说明这次行驶轨迹的直观信息。热图演示可以推导出驾驶轨迹的特定类型。

图 6: 驾驶员行为可视化



例如，你会在弯道时加速过猛吗，或者在直道上紧急刹车吗？这些工具和可视化技术在探测驾驶行为和模式方面发挥着重要作用。简单地说，这些工具和技术可以更好地帮助我们理解收集的数据，并提供直接的补充性意见。更容易检测到驾驶模式的异常值。

事件分析可用于将过往的风险经验与驾驶行为相关，就可以针对驾驶进行特定的事件分析（比如碰撞）。当然我们从风险的角度（即索赔数据）采用车联网技术并非必要。

推导驾驶习惯并开发相关服务

为了提供新的服务，确实有可能仅需要定义驾驶员情况或检测非典型路线。简单的 GPS 数据（可以通过智能手机或者与车辆连接的内置盒子收集数据）可以帮助了解投保人的驾驶行为。非监督分类算法已经被用来鉴别这种行为信息。

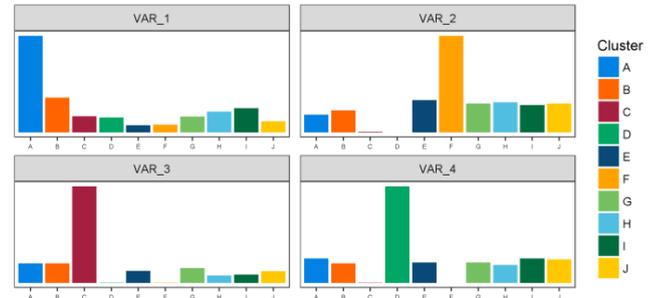
详细地说，这项研究从保险公司的业务组合中随机抽取约 1,000 辆汽车的信息，进行详尽分析：既粗略分析总体路程，也定期分析数据提供商提供的具体“事件”。对于保险公司来说，这项分析任务的难度有两层：要验证这些提供的信息质量可靠并做出解释，同时还要鉴别改变的路径。

尽管难以量化，但投资回报率确定是有的。这项研究侧重于保险公司产品条款的可能演进，持续地关注被保险人的需求和感受。

路程分析

基于我们在风险、统计和定价方面的专业技能，可以定义不同的路线类别和类型，然后对每一个类别进行分析和后验解释（参见下图）。

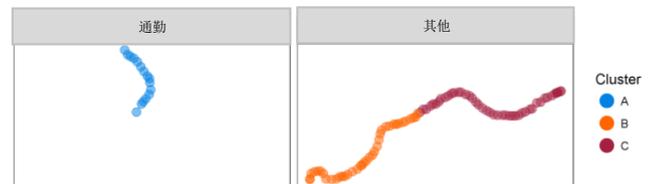
图 7: 聚类分析



基于诸如一天中的具体开车时间、距离起点的距离、或该区域的地理位置在内的大概 30 个解释变量，定义不同类型的路程。在这些路程类别的基础上，可以根据分类和分布情况来描述每名驾驶员。

在最后阶段才添加驾驶的系统性回顾及综合分析评比（PRISMA）信息。实际上，除了传统指标之外，我们还想评估非监督建模的质量。得出的结论是：对于近 60% 的驾驶员来说，他们的路程差不多也就分为三类，也就是说，聚类分析就足够稳健。

图 8: 路程分类示意图

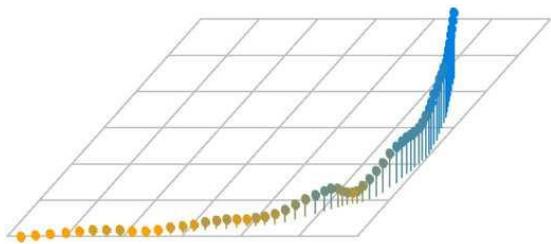


该研究表明这些分类中的路程分布体现投保人的驾驶习惯。该研究在没有事先与索赔数据关联的情况下，让保险公司更加了解他的投保人。使用相对简单的车联网数据（即低频率的原始数据，而没有使用数据提供者报告的具体事件信息）可以建立初始的驾驶情况。结合机器学习技术和业务专业知识，这些数据已经可以应用于不同目的：非典型路线和驾驶情况检测，以及地理围栏服务等。

事件分析

尽管 GPS 数据已经提供了关于驾驶概况的大量信息，车联网收集的信息可以更详细地描述在路程期间发生的特定事件。为了提供事故后救援或事故重现等个性化服务，以便更快速地进行理赔处理（或者可能是欺诈行为检测），我们可以进一步分析更精细的数据，例如加速计的信息。

图 9：加速度数据的简化轨迹重构



我们对高频率获取的数据进行研究。使用初始位置和速度数据，以及简化的物理方法可以在获取事故发生时车辆轨迹的实时三维背景信息（图 9）。

需要指出的是这些只是近似值，实际上根据变化情况推导位置是一项更复杂的任务。但通过这些粗略的视觉图像，保险公司已经能够更清楚地了解事件并定义自己的指标。

特定情境下事件的质量控制至关重要，比如利用车联网提供的事故信息适时安排救援。事故检测如果出错，除了影响公司品牌形象之外，还可能导致启动不必要的救援过程并发生费用，影响公司净利润。如何通过补充性观点深入了解和理解数据显而易见与投资回报率（ROI）有关。

因此理解车联网数据是所有创新服务的核心。即便样本量较小的情况，分析结论仍然合理，也为未来大型项目铺平了道路。

大数据与分析学不仅仅是用最新算法分析海量数据，更是一门如何将项目由小到大拓展成有高附加值项目的一门学问。

车联网发展前路不会是一番坦途：**汽车保险行业将有一个真正的机会来重塑自己、改变面貌、避免漂移。定价之外，预防驾驶风险和提升客户体验都至关重要。**

Milliman

明德是世界上最大的保险精算及相关产品和服务供应商之一。该公司为财产保险、人寿保险和金融服务、医疗保健和员工福利等领域提供咨询服务。明德成立于 1947 年，是一家独立公司，在全球主要城市设有办事处。



联系方式

蒋冠军

合伙人和精算顾问，上海
guanjun.jiang@milliman.com

Remi BELLINA, IA

精算顾问和数据科学家，巴黎
remi.bellina@milliman.com

Antoine LY, IA

精算顾问和数据科学家，巴黎
antoine.ly@milliman.com

Fabrice TAILLIEU, IA

合伙人和精算顾问，巴黎
fabrice.taillieu@milliman.com

milliman.com

© 2018 Milliman, Inc. 版权所有。本文的内容代表作者的观点，但不代表明德公司的观点。明德没有对本文信息进行认证，也不保证此类信息的准确性和完整性。使用这些信息是自愿的，除非已经对其准确性和完整性进行了独立评估，否则不应作为依据。未经明德明确同意，不得复制本文。