主动防御的双结构网络

尹浩^{1*}, 郭东超^{1*}, 吕勇强¹, 杨鹏^{2,3}, 赵志为⁴, 张尧学⁵

1. 清华大学信息技术研究院, 北京 100084

2. 东南大学未来网络研究中心, 南京 211189

3. 计算机网络和信息集成教育部重点实验室 (东南大学), 南京 211189

4. 电子科技大学计算机科学与工程学院, 成都 611731

5. 清华大学计算机科学与技术系, 北京 100084

* 通信作者. E-mail: h-yin@mail.tsinghua.edu.cn, dongchaoguo@mail.tsinghua.edu.cn

收稿日期: 2017-08-10; 接受日期: 2018-01-08; 网络出版日期: 2018-11-27

国家重点研发计划 (批准号: 2016YFB1000102) 和国家自然科学基金 (批准号: 61672318, 61631013) 资助项目

摘要 网络空间安全攸关人类福祉与国家利益, 但当前网络空间安全存在重大挑战: 互联网体系结构具有优越的互联互通性和开放性, 但也导致其安全问题无法彻底解决, 处处设防却处处难防; 传统安全防御机制依赖安全威胁的先验信息, 难以应对不断演化的安全威胁. 为了应对上述挑战, 本文提出了一种新型网络空间安全防御理论体系: 构建与互联网主结构平行的“动态、异构、冗余”的基于播存思想的次结构网络, 辅助现有互联网体系结构而形成双结构网络; 提出不依赖安全威胁先验信息的数据与知识联合驱动的新型主动防御机理; 提出双网透明接入与前置主动防御技术等关键支撑技术. 为改变我国网络空间安全领域受制于人、被动防御的现状, 本文试图从演进式改变网络体系结构角度提出完备解决方案.

关键词 双结构网络, 主动防御, 统一内容标签, 透明接入, 大数据

1 引言

承载着人类各种信息活动的人造虚拟空间——网络空间, 起源于机器之间互连网络 ARPANET, TCP/IP 等协议的引入则促使其转变为一个具有良好开放性和扩展性的网络与网络互联系统, 随着近年来信息化快速发展更是转变为基于互联网、电信网等信息技术基础设施、互联人机、承载人类信息活动的空间. 网络空间对现实物理空间的影响日益深远, 维护网络空间安全 (cyber security) 已经成为各国生存和发展的核心要素. 然而当今互联网体系结构固有缺陷以及安全威胁手段不断变化, 给网络安全防御方法的研究提出了挑战.

引用格式: 尹浩, 郭东超, 吕勇强, 等. 主动防御的双结构网络. 中国科学: 信息科学, 2018, 48: 1651–1669, doi: 10.1360/N112017-00171
Yin H, Guo D C, Lyu Y Q, et al. Dual-structural network of active defense (in Chinese). Sci Sin Inform, 2018, 48: 1651–1669, doi: 10.1360/N112017-00171

(1) **单一体系结构难以兼顾互联互通与安全可控**. 在体系结构层面, 基于端到端通信的互联网体系结构具有优越的互联互通性和开放性, 但其很难调和互联互通与安全可控的对立矛盾, 其根本原因是二者在互联网体系结构上产生了紧耦合, 难以同时满足. 网络空间组成构件自身存在未知漏洞、安全威胁来源和路径海量难控^[1], 导致处处设防却依然处处难防. 从网络体系结构的顶层视角出发, 基于解耦分治的思路, 研究兼顾互联互通与安全可控的创新性体系结构, 是网络空间安全研究迫切需要解决的难题.

(2) **先验信息驱动的安全防御方法难以有效应对日趋复杂多变的安全挑战**. 在机制与方法层面, 常规安全防御机制依赖于安全威胁的先验信息, 基于先验信息驱动的方法建立防御手段和安全认知, 其防御效果和安全分辨能力受限于人类的预认知水平, 应对变化的能力差. 如何借助新型网络体系结构, 研究数据驱动知识产生以补充先验知识不足, 构建数据与知识联合驱动的创新防御机制和安全认知维度, 以有效应对复杂多变的网络安全挑战, 是网络空间安全研究中亟需突破的难题.

对于网络空间这个由多种异构系统组成的“系统的系统”, 想要在当前单一的互联网体系结构框架下, 以较低成本保障互联互通性、安全性是极为困难的, 这也促使网络安全领域研究者和实践者思考如何从体系结构和方法论角度提出具有内生安全机制的主动防御网络架构. 在体系结构上, 研究者多采用彻底解决互联网互联互通与安全可控间矛盾的思路, 如命名数据网络^[2]、基于角色的体系结构^[3]、可信网络体系结构^[4]等, 试图以“重构”的方式解决网络安全与规模问题^[5], 但由于全网部署代价过于昂贵, 研究进展缓慢; 在防御方法上, 传统防御方法依赖先验知识、被动应对安全威胁, 近年来研究人员提出了许多新的防御方法, 如美国的移动目标防御 (moving target defense, MTD)^[6]和定制可信空间 (tailored trustworthy spaces, TTS)^[7]、解放军信息工程大学邬江兴院士的拟态防御^[8]、北京邮电大学方滨兴院士的使命确保技术等, 其共性特点是“以变制变”, 靠“动态、异构、冗余”的方法改变游戏规则, 扭转防御态势, 因此可以被看作是“主动防御”的方法. 与传统防御方法相比, 主动防御效果提升明显, 具有非常好的研究价值和应用前景. 除了在体系结构和防御方法层面的安全研究外, 基于共性基础安全技术研究设备、传输信道和内容等方面的先进安全技术, 如可重构硬件^[9]、物理指纹^[10]、加密、量子通信^[11]和网络态势感知^[12,13]等, 都是本领域的共性关键基础技术和研究热点.

本文的研究目标为从体系结构和方法论层面入手, 提出具有主动防御能力的网络体系结构及与之配套的防御方法, 应对网络空间安全挑战. 为应对网络空间安全挑战, 本文提出主动防御的双结构网络 (如图 1 所示), 其解决思路及其设计思想为: (1) 在体系结构层面, 借助生物学中的共栖概念, 为当今互联网体系结构设计一个“冗余、异构”的辅助的次结构, 构成主动防御的双结构网络基础设施. (2) 在防御方法层面, 在双结构网络基础上, 采用网络安全的第 4 范式^[14,15], 提出数据与知识联合驱动的主动防御技术, 不依赖于网络安全威胁的先验知识也能以变制变地应对潜在的安全威胁, 实现了防御方法层面的动态性.

在体系结构层面. 从网络体系结构互联互通与安全可控之间的对立要素分析出发, 提出主动防御的双结构网络模型. 在主结构互联网的基础上, 并行构建异构、动态、冗余的次结构网络^[16,17], 即面向用户的内容驱动网络 (user-oriented content driven network, UCDN), 脱离地址驱动机制, 转为统一内容标签 (uniform content label, UCL)^[18]驱动, 变点到点多跳互通结构为类星型广播结构. 次结构广播分发包含高价值密度的 UCL 而广播不限共享人数, 内容则由主结构分发至前置接入设备, 通过 UCL 辅助提高内容安全性的同时, 不降低现有网络服务质量. 双结构共栖共存, 主结构有毒带菌, 次结构安全可信, 改变游戏规则, 开辟第二战场. 实际上, 采用双网隔离、网闸互通的双网络安全结构应用已久, 但本文提出的双结构网络解决了传统的双网络的 3 个主要不足, 即 (1) 双网同构问题: 同构网络间安全威胁可复制, 网闸一开, 毒菌全来; (2) 复杂和脆弱性问题: 传统双网复杂性类似, 冗余网络很

图 1 (网络版彩图) 双结构网络模型

Figure 1 (Color online) Dual-structural network model

难做到完全自主可控, 其受制于软硬件供应链的安全脆弱性依然存在; (3) 网络孤岛: 由于用户缺乏网络透明接入, 隔离导致冗余网络成为孤岛, 使得冗余网络对第一网络难以构成有效的安全支持。

在防御方法层面, 大数据技术的兴起使得网络安全技术的发展进入了以数据密集计算为典型特征的“第 4 范式”阶段^[14]。基于双结构网络模型, 本文提出数据与知识联合驱动的双结构网络主动防御方法, 通过数据与知识联合驱动: (1) 实现双结构网络内容交换中心动态可重构, 同时借助统一内容标签 UCL 技术推动计算模式从数据为中心向内容为中心转变; (2) 通过能级模型的内容跃迁管控与类星广播模式解决次结构安全可信问题; (3) 通过互联网内容与次结构能级知识在用户端的牵连关系, 推动互联网产业生态内生安全激励机制的建立。数据与知识联合驱动的主动防御方法有如下优点: (1) 数据与知识联合驱动双结构网络以变制变, 保证次结构安全可信; (2) 同时变中亦有不变, 双结构网络的安全价值导向不变, 次结构网络广播 UCL 标示互联网内容的安全能级, 通过 UCL 顺应用户的安全价值取向强关联网络服务提供商 ISP (Internet service provider)、内容提供商 ICP (Internet content provider) 的网络安全与经济效益, 激励其增加安全投入, 激发互联网的自生“免疫力”, 建立内生安全网络生态。

本文内容组织如下: 本节指出了当前网络空间安全面临的挑战和亟需解决的难题, 并从体系结构、方法论等角度综述了近年来研究人员提出的具有主动防御能力的体系结构和方法。相比于传统防御方法被动应对网络空间安全威胁, 所谓的主动防御依靠在网络空间各个部分各个层次构造实现“动态、异构、冗余”, 使得攻击者难以轻易掌握被攻击者的系统特征, 而被攻击者则依靠大数据和先验知识联合驱动的防御方法识别潜在的安全威胁。然后概述了本文应对挑战的解决思路、系统设计思想。本文整体研究以第 2、3 节的理论与方法为基础, 其中第 2 节详细介绍主动防御的双结构网络理论模型与体系结构, 而第 3 节提出构造数据与知识联合驱动的主动防御机理及方法。基于上两节所提出的理论和方法, 提出双结构网络内容交换 (第 4 节) 与前置防御 (第 5 节) 两大关键技术及原型, 其中

第 4 节阐述双结构网络内容交换技术研发与原型构建, 第 5 节介绍双结构网络用户透明接入与前置主动防御. 第 6 节则介绍了双结构网络未来可能的典型应用场景, 最后第 7 节给出了本文的结论.

2 主动防御的双结构网络理论模型与体系结构

本节阐述双结构网络设计原理及体系结构, 从解决“单一体系结构难以兼顾互联互通与安全可控”这一挑战入手, 基于复杂网络理论和泛函形式化描述方法, 量化分析互联网联通性与安全性要素, 并基于此研究互联网单一体系结构解耦方法, 建立主动防御的双结构网络模型与体系结构, 面向用户设计双结构网络组成构件、协议架构、通信模式, 以及运行机制, 提出双结构网络安全量化评估模型, 构建双结构网络的理论基础. 作为补充, 本节最后从性能(能量消耗、互联互通)的角度对复杂单一网络理论极限进行了分析, 以此来进一步阐述单一结构网络难以达到通信网性能理论极限, 论证了体系结构变革的必要性.

2.1 互联网安全性与联通性要素分析和解耦

网络安全性与联通性的关联要素组成复杂, 需要按照一定的指标体系或者维度组织. 本小节基于复杂网络理论^[19], 从拓扑维度出发, 结合常见的安全性指标体系与经典安全理论中的疾病传播模型^[20, 21], 对安全性和联通性要素进行量化建模, 基此分析单一网络体系结构联通性与安全性的解耦方法, 指导双结构网络模型与体系结构研究.

安全性与联通性要素. 互联网安全性相关的要素很多, 对应的安全性维度、安全指标体系存在较大差异, 统一建模分析难度很大. 本小节聚焦在体系结构层面紧耦合、强对立的安全性及联通性这一对矛盾关联的主要要素, 考虑常见的几类安全性关联指标, 包括可靠性、一致性、私密性、可信性, 以及其他安全要素等, 主要从体系结构安全性维度考察安全性与联通性要素. 以基于疾病传播模型考察拓扑层面的安全关联要素为例, 包括网络拓扑邻接关系 $A(t)$ 、节点安全性 $V(t)$ 、节点感染或攻击概率 I 、自感染概率 E 、自愈概率 Δ 、网络联通性 C 等细分要素; 收集、整理和研究主要关联性要素, 通过量化分析筛选与网络体系结构安全性强关联的要素, 完成建模和分析.

安全性形式化描述. 基于上述要素, 根据网络传播动力学理论, 网络体系结构层面的安全性变化的过程可表示为如下的泛函数 (以其中部分要素为例):

$$\frac{d}{dt}V(t) = f_1(V, C, I, \Delta, E, t), \quad (1)$$

基于式 (1) 可量化分析安全性及关联要素之间的紧耦合关系, 分析过程采用泛函分析方法结合模型辨识, 考虑到安全要素的复杂性, 需考虑定性和定量分析结合. 为兼顾安全性和联通性对立要素, 对于网络体系结构改变所需成本可表示为

$$J = \int_0^T f_2(C, I, \Delta, E, S_1, S_2, t) dt, \quad (2)$$

其中 S_1 和 S_2 分别表示建设改造成本和系统的信息共享成本. 可通过设计新型网络模型和体系结构, 对安全关联要素进行调节, 使得最大化系统稳态安全性指数 $\|V_\infty\|$ 的同时兼顾代价 J 和联通性 C , 即

$$\max \|V_\infty\| \quad \text{s.t.} \quad C \geq C_{\min}, J \leq J_{\max}. \quad (3)$$

图 2 (网络版彩图) 双结构网络体系结构

Figure 2 (Color online) Dual-structural network architecture

由此将设计新型网络模型和体系结构的问题转化为最优化问题, 常采用的解决方法包括但不限于静态优化, 如果考虑参数的动态性则解决方法包括随机优化等.

双结构安全性分析. 本小节提出的双结构网络实际上是采用异构双结构组成的超网络 (network of networks, NoN) [22, 23], 利用不同结构优势互补, 克服单结构网络的缺陷. 基于拓扑分析理论, 两种异构网络组成的超网络邻接矩阵如下所示:

$$A = \begin{bmatrix} (A_1)_{n \times n} & (A_0)_{n \times m} \\ (A_0^T)_{m \times n} & (A_2)_{m \times m} \end{bmatrix}, \quad (4)$$

其中主结构 A_1 代表当今互联网, A_2 为次结构, 两网络通过 $(A_0)_{n \times m}$ 指定模式连接. 为最大化超网络的稳态安全性指数 $\|V_\infty\|$, 次结构模型设计为简单的近似无标度网络的类星结构, 以疾病传播模型安全性评价为例, 次结构可采用简单的网络组件实现, 因而节点自感染率 E 低、自愈速率 Δ 高, 易达到高稳态安全性; 次结构无需处处设防, 只要对类星结构中关键节点以及主次结构之间的通道进行重点防范, 即可降低主次结构之间的感染速率 I . 此外, 从网络拓扑理论分析, 近似无标度网络的类星型次结构网络平均路长远低于主结构, 使得信息共享成本 S_1 (单位 bit 的共享成本) 极大降低, 实现比主结构更高的共享能效, 实际降低了次结构的运营成本.

2.2 主动防御的双结构网络模型和体系结构

基于 2.1 小节中安全性与联通性的解耦分析, 本小节提出以优化安全性为目标的主动防御的双结构网络体系结构, 建立双结构网络组成构件、协议架构、通信模式, 以及运行机制, 如图 2 所示.

双结构网络组成构件. 双结构网络主要构件包括主结构互联网、基此平行构建的次结构网络, 以及双网互通、牵连构件、基础支撑性构件等 4 大部分. 双网互通构件即双网内容交换中心, 实现双网内容交换; 双网牵连构件即双网透明接入设备, 支持用户透明访问双结构网络并实现前置主动防御; 基础性构件即双网信息平面与控制平面, 信息平面用于支撑网络大数据相关的处理, 控制平面用于驱动次结构网络, 以及互通牵连构件的可控动态变化, 提升主动防御能力.

次结构体系结构. 与面向端到端通信的主结构不同, 次结构设计为近似无标度的类星型单向广播网络^[17], 次结构中心库 (UCDN 中心库) 内容可以通过卫星、主干光纤等多种方式进行广播, 结构简单、易于完全自主产业链实现, 可实现自主可控, 极大降低了网络中节点自发感染速率 E . 次结构只有单向下行的广播通道, 所以用户对次结构中心库不具备传染能力, 感染速率 I 为 0.

双结构交互机制. 双结构网络之间的交互机制是双结构网络的核心运行机制, 其互通交互机制通过双网内容交换中心实现, 其隔离牵连机制通过双网透明接入设备实现. 双结构网络内容交换中心建立安全能级调制的内容交换通道, 采用异构动态的交换架构、机制和协议, 阻断主结构网络安全威胁基于传统互联网架构和协议的渗透通路, 保证 UCL 在主次结构间的安全生成和传递. 双网透明接入设备支持用户在不改变传统上网模式的方式下安全接入双结构网络, 双网相互隔离, 通过在设备上的基于 UCL 的前置内容计算和主动防御实现主次结构安全牵连.

2.3 网络安全性评价模型

网络安全相关指标包括网络可靠性、一致性、私密性、可信性和安全性等, 各项指标又可以对应不同的安全维度, 如拓扑结构安全性、协议安全性、应用安全性等, 建立可全面涵盖网络安全问题的安全性评价模型具有极高的难度. 从泛函分析出发, 定义广义网络安全性评价模型为如下形式:

$$S = F(S_R, S_C, S_P, S_T, S_O), \quad (5)$$

其中涉及的安全关联指标包括可靠性 (S_R)、一致性 (S_C)、私密性 (S_P)、可信性 (S_T), 以及其他安全要素 (S_O). 基于模型辨识和泛函分析, 同时考虑一定的模型降阶技术, 简化该安全性描述, 建立可量化计算的安全性评价模型. 假设以网络拓扑维度的疾病传播安全性作为其他安全要素 S_O , 如采用简单的线性模型结构, 则基于函数式 (1) 可将拓扑维度的安全性表示为系统稳态安全, 形式如下:

$$\|V_\infty\| = c_0\|V\| + c_1\|C\| + c_2\|I\| + c_3\|\Delta\| + c_4\|E\|. \quad (6)$$

对给定网络 G 及式 (6), 在满足式 (3) 中约束条件情况下, 能达到的稳态安全性指数 $\|V_\infty\|$ 越高说明系统安全性越高. 可通过双网信息平面收集的网络安全性测量数据, 对式 (6) 安全性评价模型进行辨识, 同时结合面向网络用户直观易懂的安全分级描述方法, 制定面向安全能级定义的连续指标离散化策略, 提高对双结构网络安全性的评价能力.

2.4 复杂网络系统的特点及性能极限

本小节从信息共享能量消耗、互联互通性两个视角出发研究, 论证单一结构网络难以达到相应理论极限, 从而得出双结构网络才能达到理论极限的结论.

首先介绍了理想信道传输信息所消耗能量的极限问题, 以此说明当今互联网在信息共享方面的低效性. 然后, 从网络平均最短路径长度这一影响网络信息分发的统计参量入手, 论述了拓扑结构为无标度网络或类星型的次结构的信息分发的高效性. 最后, 结合上述论述, 得出双结构网络比单一结构互联网更加逼近理论极限的结论.

2.4.1 网络的信息共享能耗与效率

本小节复述文献 [24] 的推论, 从信息传输所耗费能量与信息传输效率问题入手, 分析包括当今互联网在内的任意信息系统在信息共享能耗方面的理论极限.

首先需要引入 Shannon-Hartley 定理, 其衡量了在给定加性 Gauss 白噪声 N 的条件下, 给定限频 B 信道 C 在观察时间 T 内的信息传输速率理论上限 C_U . 假设输入信号为 S , 输入信号方差 P_{S_0} , 噪声信号 N 的均值为 0, 功率谱密度在 $[-B, B]$ 之间为常数 n_0 , 平均功率为 $P_{N_0} = n_0 B$, 则 C_U 可表示为:

$$C_U = B \log_2 \left(1 + \frac{P_{S_0}}{P_{N_0}} \right), \quad (7)$$

C_U 的单位为 bits/s. 可以采用另一种方式定义输入信号平均功率 P_{S_0} , 如下所示:

$$P_{S_0} = \frac{C_U E_b}{B} \Rightarrow \frac{C_U}{B} = \frac{P_{S_0}}{E_b}, \quad (8)$$

其中 E_b 代表传输每 bit 信息所消耗的能量. 如果只考虑热噪声, 即考虑给定通信系统性能的上限, 令 K 为 Boltzmann 常数, 则

$$n_0 = KT, \quad (9)$$

将式 (9) 代入式 (7), 可得到如下公式:

$$\frac{C_U}{B} = \frac{\ln \left(1 + \frac{P_{S_0}}{KT B} \right)}{\ln 2}. \quad (10)$$

当带宽 B 无穷大时, 式 (10) 的极限为

$$\frac{C_U}{B} = \frac{P_{S_0}}{KT \ln 2}. \quad (11)$$

联立式 (8) 和 (11), 可得

$$E_b = (\ln 2) KT. \quad (12)$$

已知 Boltzmann 常数 $K \approx 1.3806 \times 10^{-23}$, 考虑常温 27°C (即 $T = 300\text{K}$), 假设全球有 60 亿人, 每人每天需要的信息约 $1 \text{ GB} \approx 8 \times 10^9 \text{ bit}$, 则传输这些信息最少消耗能量 (单位: J) 为

$$E_b \times (6 \times 10^{10}) \times (8 \times 10^9) = 0.6931 \times 1.3806 \times 10^{-23} \times 300 \times (6 \times 10^{10}) \times (8 \times 10^9) \approx 1.38. \quad (13)$$

由式 (13) 可见, 理论上传输全球每人每天所需信息所需能量消耗极其微小. 然而, 信息共享的理论极限却与当今互联网体系结构下的信息共享实际情况之间存在巨大落差. 造成这种落差的原因在于互联网采用端到端通信、共享通信链路带宽的方法共享信息, 故共享人数受限于通信链路的带宽. 当今互联网面临着流量爆炸性增长、服务质量保证、节能等问题, 而其端到端设计原则虽保证了互联互通应用但对于信息共享应用支持不足. 此外, 互联网的复杂路由机制造成路由器能耗巨大. 种种原因使得互联网体系结构下的规模化的信息共享效能低下, 故在当今互联网单一体系结构下无法达到高效的信息共享.

2.4.2 最少跳数内容分发网络

互联网本身并不是支持内容共享服务的最佳体系结构, 面向端到端通信的体系架构失配于内容共享类服务 (应用) [17]. 最初的互联网只应用于文件传输、电子邮件等“端到端”的信息交换服务, 节点间以大致均衡的方式连接, 任意节点度 k 分布 $P(k)$ 可以利用 Poisson 分布描述:

$$P(k) = e^{-c} \frac{c^k}{k!}, \quad (14)$$

其中存在一个平均值 c 作为特征标度, 即 Poisson 度分布网络可以由其平均值表征. 随着 Web 等应用的出现, “多对多”的内容共享服务逐步占据主流, 超过信息交换服务, 多数用户只享用少数信息资源, 符合“二八定律”. 互联网应用组成的网络由服从 Poisson 分布的随机网络为主, 逐步演化成服从幂律分布的无标度网络和随机网络的混合^[17].

无标度网络任意两点间平均最短路径长度极短, 从而在信息分发效率上具有显著优势. 下面分析无标度网络平均路长以及理论上实现高效信息分发的可行性. 幂律形式度分布 $P(k)$ 如下所示:

$$P(k) \sim k^{-\alpha}. \quad (15)$$

对于度分布服从式 (15) 的由 N 个节点组成的有限大小无标度网络, 根据幂指数 α 不同平均最短路径长度 $\bar{L}$ 也不同,

$$\bar{L}^{\alpha>3} = \frac{\ln N + \ln(\frac{\alpha-1}{\alpha-3}) - \frac{2}{\alpha-1} - \gamma}{\ln(\frac{\alpha-2}{\alpha-3}) + \ln m} + \frac{1}{2}, \quad \alpha > 3; \quad (16)$$

$$\bar{L}^{\alpha=3} = \frac{\ln N - \ln(\frac{m}{2}) - 1 - \gamma}{\ln \ln N + \ln(\frac{m}{2})} + \frac{3}{2}, \quad \alpha = 3; \quad (17)$$

$$\bar{L}^{\alpha<3} = \frac{(\frac{2}{\alpha-1}) \ln N + \ln(\frac{\alpha-1}{3-\alpha}) - (\frac{2}{\alpha-1}) - \gamma}{(\frac{3-\alpha}{\alpha-1}) \ln N + \ln(\frac{\alpha-2}{3-\alpha}) + \ln m} + \frac{1}{2}, \quad 2 < \alpha < 3. \quad (18)$$

式 (16)~(18) 的结论均是对于稠密网得到的结论^[25], 其中 m 为最小度数, $\gamma \simeq 0.5772$ 为 Euler 常数. 当 N 趋于无穷大时, 式 (16)~(18) 可以简化得到如下结论:

$$\bar{L}^{\alpha>3} \sim \ln N, \quad \alpha > 3; \quad (19)$$

$$\bar{L}^{\alpha=3} \sim \frac{\ln N}{\ln \ln N}, \quad \alpha = 3; \quad (20)$$

$$\bar{L}^{\alpha<3} = \frac{2}{3-\alpha} + \frac{1}{2}, \quad 2 < \alpha < 3. \quad (21)$$

另有文献证实, 对于 $\alpha > 2$ 的情况, 下限为 $\bar{L} \sim \ln \ln N$ ^[26], 可见无标度网络平均路长介于 $\ln \ln N$ 和 $\ln N$ 之间. 由此可知, 具有幂律度分布的次结构平均路长极短, 且平均路长并不随着网络增加而显著增加, 故而在信息分发方面具有显著优势.

理论上无标度网络平均最短路径长度的下限, 工程上也能达到. 文献 [27] 指出, 当 $2 < \alpha < 3$ 之间, 采用贪婪路由策略可使得网络平均最短路径长度达到下限 $\bar{L} \sim \ln \ln N$. 根据上述分析, 假设 WWW 中包含网络节点个数 $N = 10^{10}$, 利用贪婪路由平均距离可达 $\bar{L} \sim 3$, 但是当前互联网单一体系结构上并没有对应的实现方案. 仅就数据分发效率角度来说, 双结构网络可以作为一种良好解决方案, 在终端用户实际访问数据前就采用广播方式将 UCL 发送到终端并存储, 用户直接访问终端即可得到包含高价值密度的 UCL, 从这个角度来说本文所提出的双结构网络突破了当今互联网跳数限制的性能极限.

2.4.3 网络体系结构的未来演化

从能量消耗和信息共享两个角度出发, 上文分别论述了理想网络结构能达到的理论极限, 预示着互联网体系结构的未来演化仍然有很大空间, 如图 3 所示. 单一结构网络具有的固有缺陷使其效率低下, 难以低能耗地进行高效信息分发, 因此本文提出采用双结构网络有效克服单一结构网络的缺陷.

本文所提出的次结构为具有近似幂律度分布 (无标度) 的类星型网络, 在体系结构上与内容共享类服务相配. 无标度网络任意两点间平均最短路径长度极短, 能够实现高效路由, 从而在信息分发效

图 3 (网络版彩图) 未来网络体系结构变革的驱动力与方向

Figure 3 (Color online) Driven factors and evolution of future Internet structure

率上具有显著优势. 次结构借助广播, 在空间上以较低能耗任意复制信息比特而不限制共享人数; 而借助内容存储机制, 在时间上任意复制信息比特. 信息共享类应用可从主结构互联网剥离出来由次结构承担, 而主结构则主要承担互联互通类交互式应用, 主次分工明确, 释放了主结构的带宽, 解决了主结构面临的问题. 此外, 在双结构网络中, 我们把互联网上大量的共享类内容根据用户偏好有选择性地借助次结构分发到用户前置终端处存储, 则当用户需要这些内容时, 只需要“最后一跳”即可获取到内容, 实现了“最少跳数内容分发”, 突破了当今互联网体系结构无法突破的理论极限.

3 数据与知识联合驱动的主动防御机理及方法

以第 2 节提出双结构网络作为基础设施, 本节研究大数据环境下不片面依赖先验知识或状态数据处理的创新性主动防御机理与方法, 解决先验知识驱动的安全防御方法难以有效应对复杂多变的网络安全挑战. 以具有“高价值密度”的统一内容标签 UCL 作为内容治理的纽带, 将数据驱动的传统安全防护方法升级为数据与知识联合驱动的主动防御方法. 基于安全能级模型通过次结构网络 UCDN 可信分发 UCL, 利用内容维度的价值引导, 增强安全投入与经济效益的关联性, 促进网络空间内生安全.

常规安全防御机制通常针对被保护对象进行外部加固, 或借助对安全威胁特征的先验知识进行模式检测和匹配, 具有静态、先验和反应滞后等特征. 另外, 网络空间安全问题往往与互联网内容紧密相连, 而传统互联网体系结构针对“互联互通”的需求, 遵循“以地址为中心”的理念进行设计, 网络缺乏对所承载内容的感知、缺少必要的路径安全支持, 如果仅仅在数据层面采用处理数据的常规防御方法来处理与内容相关安全问题, 必然使安全防御效果大打折扣. 为此, 本节在双结构网络支持下高度重视内容语义和知识关联, 采取“以内容为中心”和“动态、异构、冗余”的思路, 研究互联网内容统一标识技术, 构造先验知识与动态大数据卷积联合的驱动框架, 将安全防护维度从数据处理扩展到内容认知层面, 研究面向内容的双结构网络主动防御深层机理, 提出数据与知识联合驱动的创新型主动防御方法, 建立以安全能级模型为基础的主动防御机制. 本节所提出的数据与知识联合驱动的主动防御方法是与双结构网络相配套的, 在单一主结构内部难以应用这种主动防御方法. 次结构实现了一个与主结构隔离、异构的网络, 以可信广播安全地分发经过萃取后的来自主结构的知识, 而主结构因其安全威胁处处难防的困境使得在其内部难以构建可信高价值数据分发信道, 继而难以以较低成本实现数据与知识联合驱动的主动防御方法.

图 4 (网络版彩图) 知识与数据卷积联合的主动防御驱动框架
Figure 4 (Color online) Active-defence architecture driven by knowledge-data convolution

3.1 新型内容统一标识方法与标准体系

由于受互联网体系结构和传统数据处理方法的制约, 网络空间难以提供全网覆盖的路径安全支持, 导致常规传统安全防护方法难以有效应对当前网络空间安全问题.

为突破此重大缺陷, 本小节提出如下解决方案: 借助所提出双结构网络提供体系结构支持, 次结构网络 (UCDN) 构建全网可信空间; 以具有“高价值密度”的统一内容标签 UCL 为基础, 统一标识互联网内容, 以 UCL 编码技术将互联网内容结构化、矢量化表示, 支持海量泛在内容不同阶次的语义信息标引.

统一内容标签的引入, 使互联网内容可关联、可索引、可治理, 支持传统以数据为中心的方法向以内容和知识为中心的主动防御方法的跃升: 利用 UCL 所具有的价值密度高、语义信息丰富和结构精短简练等特点, 借助双结构网络将网络空间安全防护维度从数据扩展到内容; 以 UCL 为内容与知识层面主动防御基元, 将其对内容标识过程称为“贴标过程”, 执行贴标过程的设备称为贴标机, 主结构内 ICP 通过贴标机对内容贴 UCL, 继而将其传送给双网交换中心过滤, 再通过次结构网络可信分发 UCL, 以提升主结构网络 and 用户安全防护能力.

3.2 数据与知识联合驱动的主动防御方法

基于“以内容为中心”的思想, 本小节提出将传统以数据为中心的防御升级为以内容和知识为中心的主动防御, 包括如下几个方面: UCL 的多个关键语义标识域, 包括内容话题、关键词和内容实体等, 可支持网络大数据进行富语义多维全息关联; 借助大数据科学范式和知识萃取技术, 可对内容语义信息和内容间关联特质进行深度挖掘, 基于此建立面向网络内容大数据的 UCL 语义域名体系; 构造知识与动态大数据卷积联合的驱动框架, 如图 4 所示, 建立以内容和知识为中心的创新型主动防御方法; 基于大数据信息平面和深度学习神经网络 DNN, 可实现网络空间安全态势智能认知和主动防御策略主动构建的高效方法.

图 4 所示的数据与知识联合驱动的安全防御机制, 利用大数据信息面收集到的当前主结构互联网上的网络安全态势相关大数据对旧有安全威胁先验知识进行萃取 (修正), 能够获得新的安全威胁知

图 5 (网络版彩图) 双结构网络宏观安全能级及其分解

Figure 5 (Color online) Energy level and its classification in dual-structural network

识,再以新安全威胁知识对安全态势相关大数据卷积(提取)得出当前网络是否面临未知的安全威胁的结论。具体描述如下:(1)假设卷积网络已经由现有已知的安全威胁数据训练过,则其对信息面采集到的网络态势大数据进行分析,即可推断目前网络面临的安全威胁;(2)防御策略主动构建:利用历史大数据训练得到一个 DNN 深度学习神经网络,然后利用大数据信息平面对网络态势感知得到的多源多维网络态势数据(如其中用户行为、内容等)调优之前得到的 DNN,即利用新安全威胁数据修正先验知识,使得 DNN 能够识别新安全威胁特征。

通过数据与知识联合驱动防御方法,双结构网络可自主认知常规基于先验知识的安全方法难以发现的安全问题,主动构建防御策略,即基于安全态势的认知,重新构建体系结构、管控安全能级、配置安全策略。数据与知识联合防御方法对网络体系结构、能级跃迁和前置安全的驱动机制,主要体现在 3 个方面:驱动网络体系结构实现双结构网络系统的动态可重构;驱动能级跃迁来管控安全能级之间的跃迁过程,产生适合于次结构网络的可选候选内容;驱动前置安全实现对双网透明接入设备的多策略动态配置和管理。基于数据与知识联合驱动防御方法充分体现双结构网络的“动态、异构、冗余”主动防御特征,显著降低网络空间的安全风险。

3.3 基于安全能级模型的主动防御实现机制

本小节参考电子绕核运动的轨道能级分离与跃迁理论,提出安全能级的概念及划分方法。基于安全能级模型的主动防御构成及实现机制描述如下:包含可对网络基础设施、网络运行态势、风险预警信息和多源内容数据等多维度网络大数据进行大规模并发采集与分析的大数据管理功能;以双网内容交换中心作为双结构网络互联互通和安全可控的关键节点,参考电子绕核运动的轨道能级分离与跃迁理论的双结构网络安全能级模型可分离、可管控;彼此分离且逐级趋严的网络空间安全能级划分方法,如图 5 所示,按照 UCL 所涉及的对象和重要程度,可从宏观上将安全能级由低到高划分为 5 级,例如,普通 ICP 发布的一般社会新闻,初始安全能级为 2 级,新华社发布的时政要闻,初始安全能级为 4 级。

此外,按照网络安全和内容安全两个层面可将安全能级分解映射为网络安全能级和内容安全能级,为细粒度的能级调制和跃迁管理提供策略选择量化依据。基于数据与知识联合驱动方法,以第 2 节中所提出的联通度和感染概率作为网络安全能级的表征要素,以内容质量和可信度作为内容安全能级的

图 6 (网络版彩图) 双结构网络内容交换架构

Figure 6 (Color online) Content exchange architecture of dual-structural network

表征要素, 构建更精细的能级通道管控机制与算法, 即可规范和约束主结构低安全能级数据和内容向次结构高安全能级的跃迁, 将主结构的数据计算模式升级切换为次结构的内容计算模式, 既兼顾联通性又满足安全性.

本小节借助 UCL 和安全能级模型将安全防护维度从数据扩展到内容, 通过安全价值引导, 激励网络生态的内生安全.

4 双结构网络内容交换技术研发与原型构建

第 2 节主要阐述了双结构网络的理论基础, 进而介绍了其体系结构, 特别介绍了次结构网络运行机制. 本节将介绍双网内容交换中心, 它是双结构网络交互机制的主要组成部分之一. 双网内容交换通道建立安全能级调制的内容交换通道, 采用异构动态的交换架构、机制和协议, 阻断主结构网络安全威胁基于传统互联网架构和协议的渗透通路, 保证 UCL 在主次结构间的安全生成和传递. 本节研究如何实现用于双网内容交换的能级跃迁通道, 而能级跃迁的概念在第 3 节中已有阐述; 实现对互联网输入的内容进行按需求的安全能级跃迁处理, 以输出符合次结构网络安全要求的互联网内容统一标签 UCL; 实现安全可控、具备主动防御能力的双网互联互通.

双结构网络内容交换架构如图 6 所示, 本节的研究以双网内容交换能级跃迁通道构建为中心, 从支撑关键技术体系、支撑运行体系和主动防御效能 3 个层面着手, 打造双网互联互通、主动防御的内容交换中心.

4.1 双网内容交换关键技术

双网内容交换所需的支撑性关键技术主要包括: 基于大数据信息平面的互联网内容 UCL 认证技术和数据与知识联合驱动的网络安全与内容安全测量、检测与评价两大部分. UCL 认证技术解决网

络内容的信源、内容一致性、完整性和安全性等校验和认证;数据与知识联合驱动的各项安全技术解决交换通道典型安全威胁检测、内容可信度与内容质量测量与评价等,以支持安全的能级跃迁。

基于大数据信息平面的互联网内容 UCL 认证. 双网内容交换中心收集来自互联网 ICP 的网络内容和与其对应的 UCL 标签,内含基于公钥基础设施的数字签名,可供交换中心对照注册名单认证该 ICP 身份,确认内容信源,以完成用户 UCL 认证.对于通过用户 UCL 认证的内容,交换中心对其进行能级跃迁操作,生成经过认证的完整 UCL.普通用户可根据接收到的 UCL 内含的 URL 通过互联网获得相应内容.为保证用户访问的 URL 对应内容与中心认证 UCL 时对应的状态一致,避免内容被篡改或安全状态发生变化,交换中心需要定时或基于事件触发对注册的内容进行全量、增量 UCL 认证.

数据与知识联合驱动的安全检测与测量. 主要解决对来自主结构网络的安全问题的检测和测量,供能级跃迁处理过程调用,具体包括交换通道典型安全威胁检测、内容质量测量与内容可信度评价等关键技术.

交换通道典型安全威胁检测主要围绕 UCL 展开,分别针对 UCL 标签内容、UCL 标签指向的互联网内容两类从互联网输入的数据进行安全检测.对于 UCL 标签内容,采用基于 UCL 格式规范标准进行格式边界越界校验、基于虚拟执行技术进行 UCL 标签内容解析与分析,对格式不全或不符合标准格式的 UCL 内容给出安全预警.对于 UCL 对应互联网络内容,其可能引入 SQL 注入攻击、跨权访问等针对应用系统的攻击威胁,也可能引入跨站攻击和隐私获取、挂马攻击、捆绑漏洞利用等针对客户端的攻击威胁,则通过在内容交换通道中集成调用第三方工具和服务来完成检测.

网络内容质量测量包括本体价值测量,如可阅性、相关性等,外部辅助性价值测量,如转载数、分享数等.对于本体价值测量,采用深度学习方法进行摘要、关键字及话题分类等语义提取、生成内容基本要素,对于要素不全或不符合先验共识逻辑的网络内容给出相应较低的质量评价.对于外部辅助性价值测量,则可以通过统计用户评论数、大数据驱动的用户背景与该内容的相关性分析、用户评价感情色彩解读等方法侧面了解网络内容质量.关于网络内容可信度评价,考虑到网络内容具有海量异构特点,基于互联网范围的大数据(包括信源、传播链和内容质量、用户反馈、历史信息等数据)评价网络内容可信度.

4.2 软件定义的安全能级跃迁通道

双结构网络内容交换通道是保证主次结构网络互通的唯一通道,也是次结构网络对主结构网络实现主动防御的重要关卡,通过“重兵把守”加强网络安全、“动态异构”实现主动防御、能级跃迁阻断安全威胁传播通路,多方法结合保障 UCDN 的 UCL 内容安全性.

本小节提出基于软件定义架构的双网内容交换安全能级跃迁通道,通过标准模块化、虚拟化等方法集成封装各种安全检查及处理的软硬件,支持软件定义通道入口能级和目标能级、通道能级跃迁对应的安全检查和处理等基础规则,以软件定义保障灵活配置能力、高并发处理能力和异构动态变化能力,提升性能和安全性.每个通道入口对应多层面、多视角的安全能级跃迁处理模块网络视图,通过对网络中的 UCL 数据流进行牵引调度,实现基于基础构件软件定义产生多通道方案的能力.在大规模、高并发处理环境下,通过功能虚拟化和分布式软件定义能级跃迁通道,可将通用云计算资源加入能级跃迁,扩大通道的处理能力、保证可扩展性.

在通道传输的内容从初始能级到目标能级的跃迁过程中,对 UCL 标签对应的原始网络内容的安全检查即是对格式化数据的处理过程,采用特征码基础检测、静态分析检测、虚拟化动态检测,以及逐渐成熟的大数据安全分析等方法进行安全检测.作为网络内容安全检查中传统的敏感信息检查模块,

可通过基于域名、关键字正则表达式等多种组合主动审计指定的目标内容, 对内容进行深度扫描检测, 快速、准确地分析判断内容中是否含有非法敏感信息, 并支持人工辅助校正扫描结果, 从而有效防止不良信息扩散, 为追查取证提供有力支持. 对于图片、视频等复杂内容, 则可集成领域内领先的商业产品, 封装为软件定义模块加入内容交换通道中.

综上, 由于双结构网络内容交换对象仅为格式化的 UCL 标签, 其安全防御范围相比主结构网络大大缩小, 其安全性检查更加容易和可行.

4.3 基于安全态势感知的动态能级跃迁策略优化

主结构互联网有毒带菌、开放互联、安全态势变化剧烈, 影响着双网内容交换能级跃迁管控的外因. 本小节研究网络安全态势感知驱动的能级跃迁策略动态优化机制, 兼顾整体安全态势和具体信源的安全态势, 实现动态能级与跃迁策略的调整, 提升主动防御能力.

安全态势感知的触角延伸方式包括对互联网整体安全态势的感知和具体信源的安全态势感知两种, 前者需要通过搜集互联网上公开的各种漏洞、安全通告信息等构建威胁情报库, 并根据该情报库的实时更新内容分析当前安全发展态势, 最终形成安全态势的跟踪能力. 后者的实现则需要以对指定的信源 Web 层面的资产发现为基础, 通过采集信源所使用的网站和主机的指纹、发布的内容, 结合之前积累的威胁情报库进行综合匹配, 用以判断指定信源的安全等级.

对互联网整体安全态势的感知需要对公开数据进行采集, 不断积累数据建立完善的威胁情报库, 同时保证对实时数据的快速跟踪能力. 对具体信源的安全态势感知, 需要全时扫描, 基于从情报库大数据萃取的安全威胁信息和漏洞信息, 对指定的信源进行无缝全自动实时监测, 第一时间发现信源存在的漏洞安全隐患.

5 双结构网络用户透明接入与前置主动防御

双结构网络之间的交互机制是双结构网络的核心运行机制, 其互通交互机制通过第 4 节介绍的双网内容交换中心实现, 其隔离牵连机制则通过本节提出的双网透明接入设备实现. 网络用户接入是保证双结构网络可用性和实现其安全价值引导的重要环节, 本节提出双结构网络用户透明接入机制和前置主动防御技术, 将网络用户透明、安全地接入到双结构网络, 在用户端实现经济效益与安全价值的牵连, 激发内生安全.

5.1 双结构网络透明接入机制

基于透明计算的方法, 建立由控制平面驱动的透明接入架构. 如图 7 所示, 包含接收次结构统一内容标签的 UCL 缓存模块、与主结构网络连通的互联网访问模块和前置安全认证模块, UCL 缓存和互联网访问模块硬件隔离, 互不干扰. 前置安全认证模块结合 UCL 对互联网访问模块的上行、下行数据进行完整性、一致性认证, 建立主次结构网络数据的牵连. 通过结合不同接入设备硬件体系结构和接口的共性特征, 将上述核心模块按照透明计算方法抽象成不依赖于特定硬件环境和操作系统的通用网络接入模块, 其软硬件运行方式由控制平面进行实时定制化配置. 控制平面基于数据知识联合驱动产生网络配置、安全防御等策略, 通过安全可信的次结构网络将配置信令传送到接入设备的 UCL 缓存, 接入设备按照透明计算的方式进行动态更新. 上述整个网络接入过程无需用户参与, 能够在不影响用户上网方式的前提下将用户透明、安全地接入到双结构网络当中.

图 7 (网络版彩图) 双结构网络透明接入与前置主动防御研究内容
Figure 7 (Color online) Transparent access and prepositive active defence

5.2 前置主动防御技术

双结构网络的前置主动防御主要包括数据安全和体系结构安全两方面。

前置接入设备的数据安全主要包括上行、下行数据安全, 以及 UCL 缓存安全。主结构互联网数据存在较高安全隐患, 故前置接入设备对互联网访问模块的上、下行数据, 尤其是没有本地 UCL 认证的互联网内容, 利用其 URL 中包含的域名、路径、访问文件类型等信息, 结合信息平面进行安全评价, 并基于此完成实时数据流过滤, 保障数据安全。用户端与前置接入设备之间的双向互通导致 UCL 缓存存在被用户端恶意程序篡改的隐患, 可采用 UCL 防篡改指纹编码机制及其校验方法, 对前置接入设备的 UCL 缓存进行定期校验, 及时发现 UCL 篡改并对其进行快速纠正。为降低 UCL 校验过程带来的网络访问延迟, 进一步完善内容自适应的 UCL 校验机制, 根据接入设备流量及访问内容自适应地调整编码能效, 在保障 UCL 安全的同时减少校验延迟。同时, 考虑到 UCL 模块缺少上行通道, 为快速纠正本地 UCL 缓存、恢复被篡改的 UCL, 可通过主结构网络与双网内容交换中心取得通信, 并通过动态重构软硬件运行体系防止 UCL 短期内再次被篡改。

在体系结构安全方面, 通过构建动态可重构的体系结构增强系统动态性, 进而提高系统主动防御能力。基于现场可编程阵列 FPGA 的局部动态可重构技术, 将整体系统按功能或者按时序分解为不同的组合, 并根据控制平面的配置要求动态重构本地系统结构, 支持主动防御。

5.3 基于前置 UCL 认证的内生安全激发机制

次结构网络对主结构网络的内容进行萃取和能级区分, 并将这些知识存放在 UCL 中分发到每个前置接入设备。利用 UCL 对主结构网络内容进行安全监测, 防止主结构网络内容篡改, 并对低安全能

级的内容进行用户提示. 对于互联网内容的请求, 根据其 URL 检索在 UCL 缓存中是否存在对应的 UCL, 以提取 UCL 校验信息、安全能级等, 并进一步进行数据校验和内容引导. 若未命中, 则基于双网内容交换中心的数据与知识联合驱动对互联网内容进行安全评价和内容引导, 激发内生安全.

互联网内容的实际安全能级随时间动态变化, 为保障中心认证 UCL 与实际互联网内容安全能级的一致性, 可基于 UCL 对互联网内容的一致性进行验证, 包括信源一致性、内容一致性、安全一致性等. 考虑到验证过程会给网络用户带来额外的延迟开销, 可采用轻量级快速数据校验方法. 在校验算法层面, 针对不同的数据类型、不同的软硬件资源设计动态自适应的轻量级数据校验算法, 对于视频、音频等大块数据实施边传输边校验的实时校验方式; 在数据信息共享层面, 考虑到前置接入设备通常连接多个用户, 在前置接入设备的 UCL 接入模块设置专用校验缓存, 对于已经被其他用户请求过、校验通过的数据免去重复校验过程, 减少重复计算. 之后, 基于 UCL 建立可信内容提示和安全引导机制, 对于 UCL 认证不通过, 或者能级较低的网络数据通过在网络接入设备上定义的出错处理机制进行处理, 如提示、警告、拒绝访问等, 从而建立安全与效益之间的强关联, 形成安全价值导向, 进而激发互联网产业生态的内生安全.

6 未来典型应用

本节以车联网中高价值密度数据传输场景为例, 介绍双结构网络未来可能的典型应用. 如图 8 所示, 在车联网高经济密度数据传输应用的主结构中, 智能汽车的传感器采集“人-车-环境”中的多源多维信息. 随后, 这些采集到的多源多维信息通过互联网传送到位于双结构网络内容交换中心的云端服务进行信息融合, 形成驾驶意图、风险态势、环境构建等新信息. 接着, 云端服务提供的态势感知等新信息将被封装入 UCL, 通过次结构 UCDN 网络广播到其他车辆. 通过广播网络进行安全可信分发, 极大地降低多管理系统建设成本; 另一方面, 在内容接收端, 通过结构简单、成本较低的次结构网络, 减少网络建设和运营成本. UCL 充当了高经济密度数据的标准化信息集装箱作用, 因而极大提高了传输效率.

综上, 在双结构网络基础上构建了智能网联汽车集成应用环境, 集成本文各部分研究成果, 为车联网内高经济密度数据的大范围高效传输、多种信息融合提供了良好的解决方案.

7 结论

网络空间安全攸关人类福祉与国家利益, 但当前网络空间安全面临严重挑战: 具有优秀互联互通性和开放性的当前互联网体系结构处处设防仍难以彻底解决安全问题, 严重依赖安全威胁先验信息的传统安全防御机制难以应对不断演化的安全威胁, 安全投入与市场效益的弱关联导致互联网产业生态缺乏内生安全的有效激励. 本文针对这些挑战, 提出了构建与互联网主结构平行的“动态、异构、冗余”的次结构网络, 辅助主结构互联网解决安全问题; 还提出了不依赖安全威胁先验信息的数据与知识联合驱动的新型主动防御机理; 此外, 提出了双网透明接入与前置主动防御技术等关键技术; 最后, 以车联网中高价值密度数据传输场景为例, 介绍双结构网络未来可能的典型应用. 本文所提出的主动防御的双结构网络, 利用统一内容标签提取对互联网内容的安全认知, 借助次结构网络可信广播特性, 通过数据与知识、双结构网络之间以及经济效益与安全能级之间的牵连, 弥补了互联网难以兼顾互联互通与安全可靠、缺乏内生安全激励、低效应对复杂多变安全威胁的缺陷.

图 8 (网络版彩图) 智能网联汽车集成应用

Figure 8 (Color online) Application of dual-structural network to vehicle network

国内外关于网络空间安全主动防御的研究, 多从网络的不同层次实现“冗余、异构、动态”的思想, 本文所提出的双结构网络则是从体系结构层面实现. 要以较低成本构建覆盖各个层次、整个生态链的主动安全防御机制, 尚有诸多理论和技术挑战有待我们去克服.

致谢 在此感谢东南大学李幼平院士的支持与指导.

参考文献

- 1 von Solms R, van Niekerk J. From information security to cyber security. *Comput Secur*, 2013, 38: 97–102
- 2 Jacobson V, Smetters D K, Thornton J D, et al. Networking named content. In: *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, Rome, 2009
- 3 Braden R, Faber T, Handley M. From protocol stack to protocol heap. *SIGCOMM Comput Commun Rev*, 2003, 33: 17–22
- 4 Lin C, Peng X H. Research on trustworthy networks. *Chin J Comput*, 2005, 28: 751–758 [林闯, 彭雪海. 可信网络研究. *计算机学报*, 2005, 28: 751–758]
- 5 Wu J X, Lan J L, Cheng D N, et al. *Novel Network Architecture*. Beijing: Posts and Telecom Press, 2014 [邬江兴, 兰巨龙, 程东年, 等. 新型网络体系结构. 北京: 人民邮电出版社, 2014]
- 6 Jajodia S, Ghosh A K, Swarup V, et al. *Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats*. Berlin: Springer Ebooks, 2011
- 7 Sheldon F T, Vishik C. Moving toward trustworthy systems: R&D essentials. *Computer*, 2010, 43: 31–40
- 8 Wu J X. Meaning and vision of mimic computing and mimic security defense. *Telecommun Sci*, 2014, 30: 2–7 [邬江兴. 拟态计算于拟态安全防御的原意和愿景. *电信科学*, 2014, 30: 2–7]
- 9 Majzoobi M, Koushanfar F, Potkonjak M. Techniques for design and implementation of secure reconfigurable PUFs. *ACM Trans Reconfig Technol Syst*, 2009, 2: 1–33
- 10 Jiang X, Yau W Y. Fingerprint minutiae matching based on the local and global structures. In: *Proceedings of International Conference on Pattern Recognition*, Barcelona, 2000. 1038–1041
- 11 Gottesman D, Lo H K. From quantum cheating to quantum security. *Phys Today*, 2000, 53: 22–27
- 12 Endsley M R. Toward a theory of situation awareness in dynamic systems. *Hum Factors*, 1995, 37: 32–64
- 13 Bandes R, Shimeall T, Heckathorn M, et al. Using SiLK for Network Traffic Analysis, *Analyst's Handbook for SiLK Versions 3.8.3 and Later*, 2014. <http://tools.netsa.cert.org/silk/>

- 14 Li D Y, Liu C Y, D Y, et al. Artificial Intelligence With Uncertainty. Beijing: National Defense Industry Press, 2014 [李德毅, 刘常昱, 杜鹂, 等. 不确定性人工智能. 北京: 国防工业出版社, 2014]
- 15 Gollmann D. Computer security. WIREs Comput Stat, 2010, 2: 544–554
- 16 Yang P, Li Y P. On the complementary binary future Internet architecture. Complex Syst Complex Sci, 2014, 11: 53–59 [杨鹏, 李幼平. 二元互补未来互联网体系结构. 复杂系统与复杂性科学, 2014, 11: 53–59]
- 17 Yang P, Li Y P. Secondary structure of future Internet based on broadcast-storage concept. Complex Syst Complex Sci, 2015, 12: 18–22 [杨鹏, 李幼平. 基于播存思想的未来互联网次结构. 复杂系统与复杂性科学, 2015, 12: 18–22]
- 18 Li Y P, Yang P. New mechanism for sharing cultural bigdata. China Comput Soc Newsl, 2013, 9: 36–40 [李幼平, 杨鹏. 共享文化大数据的新机制. 中国计算机学会通讯, 2013, 9: 36–40]
- 19 Wang X F, Li X, Chen G R. Complex Networks Theory and Its Application. Beijing: Tsinghua University Press, 2006 [汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用. 北京: 清华大学出版社, 2006]
- 20 Van Mieghem P, Omic J, Kooij R. Virus spread in networks. IEEE/ACM Trans Netw, 2009, 17: 1–14
- 21 Van Mieghem P, Cator E. Epidemics in networks with nodal self-infection and the epidemic threshold. Phys Rev E, 2012, 86: 016116
- 22 Wang Z T, Wang Z P. Elementary study of supernetworks. Chin J Manage, 2008, 5: 1–8 [王众托, 王志平. 超网络初探. 管理学报, 2008, 5: 1–8]
- 23 Kivela M, Arenas A, Barthelemy M, et al. Multilayer networks. J Complex Netw, 2014, 2: 203–271
- 24 Xing L, Ma J G, Ma W D. Information Sharing Theory and Network Architecture. Beijing: Science Press, 2011 [邢玲, 马建国, 马卫东. 信息共享理论与网络体系结构. 北京: 科学出版社, 2011]
- 25 Fronczak A, Fronczak P, Hołyst J A. Average path length in random networks. Phys Rev E, 2004, 70: 056110
- 26 Cohen R, Havlin S. Scale-free networks are ultrasmall. Phys Rev Lett, 2003, 90: 058701
- 27 Boguñá M, Krioukov D. Navigating ultrasmall worlds in ultrashort time. Phys Rev Lett, 2009, 102: 058701

Dual-structural network of active defense

Hao YIN^{1*}, Dongchao GUO^{1*}, Yongqiang LYU¹, Peng YANG^{2,3}, Zhiwei ZHAO⁴ & Yaoyue ZHANG⁵

1. Research Institute of Information Technology (RIIT), Tsinghua University, Beijing 100084, China;
2. Future Network Research Center, Southeast University, Nanjing 211189, China;
3. Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 211189, China;
4. College of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;
5. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

* Corresponding author. E-mail: h-yin@mail.tsinghua.edu.cn, dongchaoguo@mail.tsinghua.edu.cn

Abstract Cyberspace security is vital to state interest. Recently, there are some challenges in cyber security. In architecture for instance, although protection mechanisms are introduced and applied everywhere, the modern network architecture (well connected and open) still has difficulty in completely ensuring cyber security. It is also observed that contemporary cyber security protection mechanism highly depends on the priori information of security threats and thus will hardly address unknown potential threats. In this paper, a novel cyberspace security protection framework is proposed. A dual-structural Internet scheme that integrates the current Internet architecture with a redundant secondary structure network characterized by its broad-storage scheme, heterogeneous structure, and dynamic protection mechanism is introduced. Also, a novel active defense mechanism that is knowledge-data driven and thus independent of the priori information of the security threat is proposed. Furthermore, some key techniques such as transparent access and prepositive active defense are introduced. The theoretical and technical work proposed in this paper offers a comprehensively evolutionary solution to constructing a cyberspace in which the protection mechanism is more independent and active.

Keywords dual-structural network, active defence, uniform content label, transparent access, big data

Hao YIN was born in 1974. He obtained his Ph.D. degree in information and communication engineering from Huazhong University of Science and Technology, Hubei, in 2002. He is currently a professor at the Research Institute of Information Technology, Tsinghua University. His research interests include big data, future network architecture, content distribution network, and Internet measurement.

Dongchao GUO was born in 1985. He obtained his Ph.D. degree in signal and information processing from Beijing Jiaotong University, Beijing, in 2014. He is currently a postdoctoral researcher at the Research Institute of Information Technology, Tsinghua University. His research interests include complex networks and content distribution network.

Yongqiang LYU was born in 1979. He obtained his Ph.D. degree in computer software and theory from Tsinghua University, Beijing, in 2006. He is currently an associate professor at the Research Institute of Information Technology, Tsinghua University. His research interest focuses on usability and security of human-machine fusion systems.

Peng YANG was born in 1975. He obtained his Ph.D. degree in computer application technology from Southeast University, Jiangsu, in 2006. He is currently an associate professor in the School of Computer Science and Engineering at Southeast University. His research interests focus on dual-structural Internet architecture, innovative content identification, and cyber content governance.